

SEARCH REQUEST FORM

Requestor's Name: Jeffrey E. Russel Serial Number: 08/121,713
 Date: 5-15-96 Phone: 308-3975 Art Unit: 1811

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Please do a Genesys/PIR/Swiss Prot sequence search on
 SEQ ID NOS 56, 58, 60, 62, 64, and 66.

Thank you.
 JER

STAFF USE ONLY

5-198

Date completed: 5/12/96
 Searcher: ARTS/HAH
 Terminal time: 20
 Elapsed time: ✓
 CPU time: 40
 Total time: 40
 Number of Searches: 1
 Number of Databases: 1

Search Site
☐ STIC
☒ CM-1
☐ Pre-S

Type of Search

☒ N.A. Sequence
☒ A.A. Sequence
☐ Structure
☐ Bibliographic

Vendors

☐ IG Suite
☐ STN
☐ Dialog
☐ APS
☐ Geninfo
☐ SDC
☐ DARC/Questel
☐ Other

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:22:10 1996; MasPar time 6.67
Seconds

547.908 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-56

Description: (1-441) from US08121713B.pep

Perfect Score: 3231

S e q u e n c e :

MMVLLHAVYSIVFVDVIIK.....FIADIGIGVGMPQMKKILKM 441

1

Scoring table: PAM 150
Gap 11

Searched: 70887 seqs, 8282111 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8 9:part9 10:part10 11:part11 12:part12
13:part13
14:part14

Statistics: Mean 34.639; Variance 146.216; scale 0.237

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	DB	ID	Description
1	3231	100.0	441	13	R71381	Vaccinia virus semaph
0.00e+00						

2	608	18.8	122	13	R71385	Variola major virus s
9.74e-47						
3	158	4.9	477	13	R74175	Human collapsin.
1.37e-04						
4	158	4.9	771	13	R71380	Human semaphorin III
1.37e-04						
5	150	4.6	724	13	R71383	Drosophila semaphorin
6.35e-04						
6	144	4.5	712	13	R71384	Tribolium semaphorin
1.98e-03						
7	137	4.2	730	13	R71379	Grasshopper semaphori
7.35e-03						
8	125	3.9	650	13	R71382	Drosophila semaphorin
6.68e-02						
9	100	3.1	1025	13	R70126	Serum opacity factor
5.38e+00						
10	101	3.1	2749	3	R13887	Inositol-3-phosphate
4.54e+00						
11	96	3.0	325	2	P70428	Polypeptide encoded b
1.05e+01						
12	97	3.0	1256	5	R27746	Muramidase released p
8.90e+00						
13	95	2.9	304	4	R22403	Partial sequence of N
1.24e+01						
14	90	2.8	78	7	R39225	Hydrophobic protein d
2.81e+01						
15	89	2.8	320	12	R60800	Rape abscission/dehis
3.30e+01						
16	90	2.8	482	3	R21409	NADH dehydrogenase su
2.81e+01						
17	86	2.7	74	13	R71123	SG-1.
5.33e+01						
18	86	2.7	208	13	R71127	Synthetic protein G f
5.33e+01						
19	86	2.7	562	2	P91904	Sequence of amy B hea
5.33e+01						
20	88	2.7	672	14	R73593	Cotransporter protein
3.87e+01						
21	88	2.7	700	13	R70235	P. falciparum EBL-e2.
3.87e+01						
22	86	2.7	911	3	R15355	Human erythrocyte mem
5.33e+01						
23	86	2.7	1498	14	R77085	Hamster sulphonylurea
5.33e+01						
24	86	2.7	1582	14	R77088	Hamster sulphonylurea
5.33e+01						
25	83	2.6	45	1	R02230	Sequence of the E5 on
8.53e+01						
26	83	2.6	547	7	R34668	Ap Serotype 7 60kDa T
8.53e+01						
27	84	2.6	677	8	R43652	Plant potassium chann
7.30e+01						
28	83	2.6	699	4	R22187	Sequence of luteinisi
8.53e+01						

29	85	2.6	1274	7	R34714	Bacillus subtilis srf
6.24e+01						
30	84	2.6	1498	14	R77084	Rat sulphonylurea rec
7.30e+01						
31	84	2.6	1582	14	R77087	Rat sulphonylurea rec
7.30e+01						
32	81	2.5	338	3	R21420	Streptomyces clavulig
1.16e+02						
33	82	2.5	453	3	R14645	Tea gene product (20.
9.97e+01						
34	82	2.5	453	10	R53467	Tea T-cell transmembr
9.97e+01						
35	82	2.5	501	5	R27741	Sequence transcribed
9.97e+01						
36	82	2.5	502	3	R15226	HincII restriction en
9.97e+01						
37	81	2.5	710	14	R72856	C. albicans FKS1 homo
1.16e+02						
38	81	2.5	920	2	R11993	Glutamate receptor 5-
1.16e+02						
39	82	2.5	986	9	R25141	JAK2.
9.97e+01						
40	81	2.5	1054	11	R58610	Yeast HMG-CoA reducta
1.16e+02						
41	82	2.5	1129	13	R70830	Murine JAK2 kinase.
9.97e+01						
42	81	2.5	1451	5	R27819	CCVInsavc spike prote
1.16e+02						
43	82	2.5	1684	3	R14948	Bacterial amylase A-1
9.97e+01						
44	82	2.5	2351	2	P70448	Human factor VIII:c
9.97e+01						
45	81	2.5	3898	3	R10473	Hog cholera virus gen
1.16e+02						

ALIGNMENTS

RESULT 1
 ID R71381 standard; Protein; 441 AA.
 AC R71381;
 DT 21-NOV-1995 (first entry)
 DE Vaccinia virus semaphorin IV protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
 Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding
 activity;
 KW modulation; nerve cell growth; immune response; viral
 pathogenesis;
 KW neurological disease; neuro-regeneration; oncological
 infection.
 OS Vaccinia virus.
 PN WO9507706-A.
 PD 23-MAR-1995.

PF 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 DR WPI; 95-131177/17.
 DR N-PSDB; Q87443.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2; Page 65-67; 101pp; English.
 CC The sequence of the vaccinia virus semaphorin IV protein.
 CC The gene sequence was isolated as the A39R open reading frame
 sequence
 CC from variola, based on sequence homology searches of a
 database with the
 CC grasshopper, Tribolium and Drosophila semaphorin sequences.
 The proteins
 CC encoded by the grasshopper semaphorin I (Q87441), human
 semaphorin III
 CC (Q87442), vaccinia virus semaphorin IV, Drosophila semaphorin
 I and II
 CC (Q87444-5), Tribolium semaphorin I (Q87446) or variola major
 (smallpox)
 CC virus semaphorin IV (Q87447) genes were used to generate a
 series of
 CC peptides (R70370-R70418), which retain semaphorin receptor
 binding
 CC activity. The semaphorin derived or semaphorin receptor
 derived peptides
 CC are potent modulators of nerve cell growth, immune
 responsiveness and
 CC viral pathogenesis. They can be used in diagnosis and
 treatment of
 CC neurological disease and neuro-regeneration, immune modulation
 and
 CC diagnosis and treatment of viral and oncological infection and
 diseases.
 SQ Sequence 441 AA;

Query Match 100.0%; Score 3231; DB 13; Length 441;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 441; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

D b 1
 mmvllhavysivfvdviiikvqryindiltldifylfkmipllfilfyfangiewhkfet 60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Q Y 1
 MMVLLHAVYSIVFVDVIIIKVQRYINDILTLDIFYLFKMIPLLFILFYFANGIEWHKFET 60

D b 6 1
 seeiistyllddvlytgvngavytfsnnklnktgltnnniyittsikvedadkdtlvcgtn 120

```

|||||
Q      y      6      1
SEEIISTYLLDDVLYTGAVYTFSSNNKLNKTGLTNNNYITTSIKVEDADKDTLVCGTN 120

D      b      1      2      1
ngnpkcwkidgsddpkhrgrgyapyqnskvtiishngcvlsdiniskegikrwrfdgpc 180

|||||
Q      y      1      2      1
NGNPKCWKIDGSDDPKHRGRGYAPYQNSKVTIISHNGCVLSDINISKEGIKRWRFRDGPC 180

D      b      1      8      1
gydlytadnvipkdglrgafvdkdgydkvyilftdtigskrivkipyiaqmclndeggp 240

|||||
Q      y      1      8      1
GYDLYTADNVIPKDGLRGAFVDKDGTYDKVYILFTDTIGSKRIVKIPYIAQMCLNDEGGP 240

D      b      2      4      1
sslsshrwstflkvelecdidgrsyqrqihhsrtiktdndtilyvffdspysksalctysm 300

|||||
Q      y      2      4      1
SSLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIKTDNDTILYVFFDSPYKSALCTYSM 300

D      b      3      0      1
ntikqsfstsklegytqqlpspasgiclpagkvvphttfeviekynvlddiikplsnqpi 360

|||||
Q      y      3      0      1
NTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTTFEVIEKYNVLDDIIKPLSNQPI 360

D      b      3      6      1
fegpsgvkwfdikekenehreyriyfikensiysfdtkskqtrssqvdarlfsvmvtskp 420

|||||
Q      y      3      6      1
FEGPSGVKWFDIKEKENEHREYRIYFIKENSIIYSFDTKSKQTRSSQVDARLFSVMVTSKP 420

Db      421 lfiadigigvgmpqmkkilkm 441
      |||||
Qy      421 LFIADIGIGVGMPQMKKILKM 441

```

```

RESULT      2
ID      R71385 standard; Protein; 122 AA.
AC      R71385;
DT      21-NOV-1995 (first entry)
DE      Variola major virus semaphorin IV protein.
KW      Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
Tribolium;
KW      variola major virus; smallpox; semaphorin receptor binding
activity;
KW      modulation; nerve cell growth; immune response; viral

```


|||
Q y 1 8 2
YDLYTADNVIPKDGRLRGAFVDKDGTYDKVYILFTDTIGSKRIVKIPYIAQMCLNDEGGPS 241

Db 62 slsshrwstllkvelecdidgrsysqinhsktik 95
|||||||:||||||| || ||:|
QY 242 SLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIK 275

RESULT 3

ID R74175 standard; Protein; 477 AA.
AC R74175;
DT 01-NOV-1995 (first entry)
DE Human collapsin.
KW Collapsin; antibody; therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Binding site 9..19
FT /note="antibody binding site"
FT Binding site 51..65
FT /note="antibody binding site"
PN US5416197-A.
PD 16-MAY-1995.
PF 15-OCT-1993; 136922.
PR 15-OCT-1993; US-136922.
PA (UYPE-) UNIV PENNSYLVANIA.
PI Luo Y, Raper JA;
DR WPI; 95-193478/25.
DR N-PSDB: Q92331.
PT New antibody to human collapsin - used to inhibit the activity
of
PT collapsin, to induce neurite out-growth and to treat
individuals with
PT nerve damage.
PS Claim 2; Columns 15-18; 11pp; English.
CC An antibody capable of specifically binding at least a portion
of
CC the collapsin protein can be used to purify human collapsin
and
CC to inhibit the activity of the protein. It can be used to
induce
CC neurite outgrowth by neuronal cells and to treat individuals
CC suffering from nerve damage.
SQ Sequence 477 AA;

Query Match 4.9%; Score 158; DB 13; Length 477;
Best Local Similarity 32.7%; Pred. No. 1.37e-04;
Matches 37; Conservative 29; Mismatches 33; Indels 14;
Gaps 10;

D b 9
dkvyfffrenaidgehsgkatharigqickndfgghrslv-nkwtflkarlicsvpgpn 67
||||:| : : : | | :|:| || || || :|:|||| | | :
:

CC diagnosis and treatment of viral and oncological infection and diseases.

SQ Sequence 771 AA;

Query Match 4.9%; Score 158; DB 13; Length 771;
Best Local Similarity 32.7%; Pred. No. 1.37e-04;
Matches 37; Conservative 29; Mismatches 33; Indels 14;
Gaps 10;

```
D          b          2          4          1
dkvyffffrenaidgehsgkatharigqickndfgghrslv-nkwttflkarlicsvpgpn 299
      |||::| : :| : : | |::| || || || ::|:|||| | | :
:
Q          y          2          0          8
DKVYILFTD-TIGSKRIVKIPY--IAQMCLNDEGGPSSLSSHRWSTFLKVELECDI---D 261

Db 300 gidthfdelqdvflmnfkdpknpvvygvfttssnifkgsavcmysmsdvrrvf 352
      | : : :: : :| :: ::| || | : | ||:| |||: ::| |
Qy 262 GR-S-YRQI--IHSRTIKTDNDTILY-VFFDSPYS-K-SALCTYSMNTIKQSF 307
```

RESULT 5

ID R71383 standard; Protein; 724 AA.

AC R71383;

DT 21-NOV-1995 (first entry)

DE Drosophila semaphorin II protein.

KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;

KW variola major virus; smallpox; semaphorin receptor binding activity;

KW modulation; nerve cell growth; immune response; viral pathogenesis;

KW neurological disease; neuro-regeneration; oncological infection.

OS Drosophila sp.

PN W09507706-A.

PD 23-MAR-1995.

PF 13-SEP-1994; U10151.

PR 13-SEP-1993; US-121713.

PA (REGC) UNIV CALIFORNIA.

PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;

PI O'Connor T;

DR WPI; 95-131177/17.

DR N-PSDB; Q87445.

PT New class of semaphorin peptide(s) and polypeptide(s) - are

PT potent modulators of nerve cell growth and regeneration

PS Example 2; Page 79-82; 101pp; English.

CC The sequence of the Drosophila semaphorin II protein. The gene was

CC isolated by PCR using primers based on sequence homology between the

CC complete grasshopper and partial Tribolium and Manduca (moth) semaphorin

CC gene sequences. The products gave two different sequences,

each of which
 CC was used to obtain its respective complete sequence:
 semaphorin I
 CC (Q87444) and II. The proteins encoded by the grasshopper
 semaphorin I
 CC (Q87441), human semaphorin III (Q87442), vaccinia virus
 semaphorin IV,
 CC Drosophila semaphorin I and II, Tribolium semaphorin I
 (Q87446) or
 CC variola major (smallpox) virus semaphorin IV (Q87447) genes
 were used to
 CC generate a series of peptides (R70370-R70418), which retain
 semaphorin
 CC receptor binding activity. The semaphorin derived or
 semaphorin receptor
 CC derived peptides are potent modulators of nerve cell growth,
 immune
 CC responsiveness and viral pathogenesis. They can be used in
 diagnosis and
 CC treatment of neurological disease and neuro-regeneration,
 immune
 CC modulation and diagnosis and treatment of viral and
 oncological infection
 CC and diseases.
 SQ Sequence 724 AA;

Query Match 4.6%; Score 150; DB 13; Length 724;
 Best Local Similarity 31.8%; Pred. No. 6.35e-04;
 Matches 34; Conservative 26; Mismatches 38; Indels 9;
 Gaps 8;

D	b	2	6	5
vyfffretaveyincgkavysriarvckkdvgg-knllahnwatylkarlncsisgefpf	323			
	:: : :	:: :	: : ::	:
:				
Q	y	2	1	0
VYILFTDT-IGSKRIVKIPY--IAQMCLNDEGGPSSLSSHRWSTFLKVELECDIDGR-S-	264			
Db	324	yfnei qsvyqlpsdk srf fatftt stngligns avcsfhineiqaa fn	370	
		: : : : : :	: :: : : :	
Qy	265	YRQIIHSR-TIKTDNDTILYVFFDSPYS-K-SALCTYSMNTIKQSFS	308	

RESULT 6
 ID R71384 standard; Protein; 712 AA.
 AC R71384;
 DT 21-NOV-1995 (first entry)
 DE Tribolium semaphorin I protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
 Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding
 activity;
 KW modulation; nerve cell growth; immune response; viral
 pathogenesis;

KW neurological disease; neuro-regeneration; oncological infection.

OS Tribolium sp.

PN WO9507706-A.

PD 23-MAR-1995.

PF 13-SEP-1994; U10151.

PR 13-SEP-1993; US-121713.

PA (REGC) UNIV CALIFORNIA.

PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;

PI O'Connor T;

DR WPI; 95-131177/17.

DR N-PSDB; Q87446.

PT New class of semaphorin peptide(s) and polypeptide(s) - are

PT potent modulators of nerve cell growth and regeneration

PS Example 2; Page 85-88; 101pp; English.

CC The sequence of the beetle Tribolium semaphorin I protein.

The gene was

CC isolated by PCR using Tribolium genomic DNA. The proteins encoded by the

CC grasshopper semaphorin I (Q87441), human semaphorin III (Q87442),

CC vaccinia virus semaphorin IV (Q87443, Drosophila semaphorin I and II

CC (Q87444-5), Tribolium semaphorin I or variola major (smallpox) virus

CC semaphorin IV (Q87447) genes were used to generate a series of peptides

CC (R70370-R70418), which retain semaphorin receptor binding activity. The

CC semaphorin derived or semaphorin receptor derived peptides are potent

CC modulators of nerve cell growth, immune responsiveness and viral

CC pathogenesis. They can be used in diagnosis and treatment of neurological

CC disease and neuro-regeneration, immune modulation and diagnosis and

CC treatment of viral and oncological infection and diseases.

SQ Sequence 712 AA;

Query Match 4.5%; Score 144; DB 13; Length 712;
Best Local Similarity 32.8%; Pred. No. 1.98e-03;
Matches 22; Conservative 18; Mismatches 21; Indels 6;
Gaps 5;

D	b	2	0	6
fvn-svaygdyiffyretaveymncgkviysrvarvckddkkgphq-srdwrtsflkar	263			
:	:	:	:	:
::				
Q	y	2	0	0
FVDKDGTYDK-VYILFTDT-IGSKRIVKIPY--IAQMCLNDEGGPSSLSSHRWSTFLKVE	255			

Db 264 lncsipg 270
|:| | |

Qy 256 LECDIDG 262

RESULT 7

ID R71379 standard; Protein; 730 AA.
AC R71379;
DT 21-NOV-1995 (first entry)
DE Grasshopper semaphorin I protein.
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
Tribolium;
KW variola major virus; smallpox; semaphorin receptor binding
activity;
KW modulation; nerve cell growth; immune response; viral
pathogenesis;
KW neurological disease; neuro-regeneration; oncological
infection.
OS Grasshopper sp.
PN WO9507706-A.
PD 23-MAR-1995.
PF 13-SEP-1994; U10151.
PR 13-SEP-1993; US-121713.
PA (REGC) UNIV CALIFORNIA.
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
DR WPI; 95-131177/17.
DR N-PSDB; Q87441.
PT New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration
PS Example 1; Page 68-72; 101pp; English.
CC The sequence of the grasshopper semaphorin I protein. The
proteins
CC encoded by the grasshopper semaphorin I, human semaphorin III
(Q87442),
CC vaccinia virus semaphorin IV (Q87443), Drosophila semaphorin
I and II
CC (Q87444-5), Tribolium semaphorin I (Q87446) or variola major
(smallpox)
CC virus semaphorin IV (Q87447) genes were used to generate a
series of
CC peptides (R70370-R70418), which retain semaphorin receptor
binding
CC activity. The semaphorin derived or semaphorin receptor
derived peptides
CC are potent modulators of nerve cell growth, immune
responsiveness and
CC viral pathogenesis. They can be used in diagnosis and
treatment of
CC neurological disease and neuro-regeneration, immune modulation
and
CC diagnosis and treatment of viral and oncological infection and
diseases.
SQ Sequence 730 AA;

Query Match

4.2%; Score 137; DB 13; Length 730;

Best Local Similarity 31.1%; Pred. No. 7.35e-03;
Matches 38; Conservative 29; Mismatches 40; Indels 15;
Gaps 11;

```

D          b          2          0          8
fvntmeyndfiffffretaveyincgkaiysrvarvckhdkggphq-ggdrwtsflksrl 266
      ||:      | ::::| :| :      | | :|::| :| ||| : : ||::|||
|
Q          y          2          0          0
FVDKDGTYDKVYILFTDT-IGSKRIVKIPY--IAQMCLNDEGGPSSLSSHRWSTFLKVEL 256

D          b          2          6          7
ncsvpgdypfyfneiqtstdiiegnygggqvekliygvtfttpvnsiggsavcafsmksile 326
      :| : | : | :| : | : : : | | | :| |
||:|::||::| :
Q          y          2          5          7
ECDIDGR-S-YRQIIHSRT--IKTD-NDTI--L-YVFFDSP-YSK--SALCTYSMNTIKQ 305

Db      327 sf 328
      ||
Qy      306 SF 307

```

RESULT 8
ID R71382 standard; Protein; 650 AA.
AC R71382;
DT 21-NOV-1995 (first entry)
DE Drosophila semaphorin I protein.
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
Tribolium;
KW variola major virus; smallpox; semaphorin receptor binding
activity;
KW modulation; nerve cell growth; immune response; viral
pathogenesis;
KW neurological disease; neuro-regeneration; oncological
infection.
OS Drosophila sp.
PN W09507706-A.
PD 23-MAR-1995.
PF 13-SEP-1994; U10151.
PR 13-SEP-1993; US-121713.
PA (REGC) UNIV CALIFORNIA.
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
DR WPI; 95-131177/17.
DR N-PSDB; Q87444.
PT New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration
PS Example 2; Page 74-77; 101pp; English.
CC The sequence of the Drosophila semaphorin I protein. The gene
was
CC isolated by PCR using primers based on sequence homology
between the
CC complete grasshopper and partial Tribolium and Manduca (moth)

semaphorin
 CC gene sequences. The products gave two different sequences,
 each of which
 CC was used to obtain its respective complete sequence:
 semaphorin I and II
 CC (Q87445). The proteins encoded by the grasshopper semaphorin
 I (Q87441),
 CC human semaphorin III (Q87442), vaccinia virus semaphorin IV,
 Drosophila
 CC semaphorin I and II, Tribolium semaphorin I (Q87446) or
 variola major
 CC (smallpox) virus semaphorin IV (Q87447) genes were used to
 generate a
 CC series of peptides (R70370-R70418), which retain semaphorin
 receptor
 CC binding activity. The semaphorin derived or semaphorin
 receptor derived
 CC peptides are potent modulators of nerve cell growth, immune
 CC responsiveness and viral pathogenesis. They can be used in
 diagnosis and
 CC treatment of neurological disease and neuro-regeneration,
 immune
 CC modulation and diagnosis and treatment of viral and
 oncological infection
 CC and diseases.
 SQ Sequence 650 AA;

Query Match 3.9%; Score 125; DB 13; Length 650;
 Best Local Similarity 37.3%; Pred. No. 6.68e-02;
 Matches 22; Conservative 12; Mismatches 20; Indels 5;
 Gaps 4;

D	b	1	1	5
vyffffretavefin	cgkaiysrvarvckwdkggphrfr-nrwtsflksrln	csipg-dy	171	
	:: : :	: :: :	: :: :	
Q	y	2	1	0
VYILFTDT-IGSKRIVKIPY--IAQMCLNDEGGPSSLSSHRWSTFLKVELECDIDGRSY	265			

RESULT 9
 ID R70126 standard; Protein; 1025 AA.
 AC R70126;
 DT 31-OCT-1995 (first entry)
 DE Serum opacity factor (SOF22).
 KW serum opacity factor; diagnostic; high density lipoprotein;
 HDL;
 KW apolipoprotein; coronary disease; Streptococcus pyogenes.
 OS Streptococcus pyogenes strain 22 D734.
 FH Key Location/Qualifiers
 FT Peptide 1..29
 FT /note= "signal sequence"
 FT Protein 30..1025
 FT /note= "mature SOF22"

FT Cleavage_site 29..30
 FT /note= "putative signal sequence cleavage site"
 FT Region 780..807
 FT /note= "proline-rich motif present in the majority
 FT of streptococcal surface proteins"
 FT Region 820..828
 FT /note= "proline-rich motif present in the majority
 FT of streptococcal surface proteins"
 FT Region 849..857
 FT /note= "proline-rich motif present in the majority
 FT of streptococcal surface proteins"
 FT Region 928..935
 FT /note= "proline-rich motif present in the majority
 FT of streptococcal surface proteins"
 FT Domain 820..828
 FT /label= repeat_1
 FT Domain 848..857
 FT /label= repeat_2
 FT Domain 888..896
 FT /label= repeat_3
 FT Domain 927..935
 FT /label= repeat_4
 FT Domain 990..995
 FT /note= "LPXTGX motif, found in surface proteins
 FT from gram positive bacteria, with a
 FT conservative replacement T to S"
 PN W09506721-A.
 PD 09-MAR-1995.
 PF 30-AUG-1994; U09926.
 PR 01-SEP-1993; US-115227.
 PR 22-AUG-1994; US-115227.
 PA (UYRQ) UNIV ROCKEFELLER.
 PI Fischetti VA, Rakonjac J, Robbins J;
 DR WPI; 95-115437/15.
 DR N-PSDB; Q83181.
 PT A process for producing a polypeptide with apolipoproteinase
 PT activity - by cloning a DNA sequence coding for a protein
 with
 PT serum opacity factor activity and isolating said polypeptide.
 PS Claim 8; Fig 4A-C; 46pp; English.
 CC SOF22 (serum opacity factor 22) is encoded by the sof22 gene
 (Q83181)
 CC from Streptococcus strain D734 and flanking regions. SOF22 has
 CC apolipoproteinase activity, cleaving apolipoprotein APO-1,
 from high
 CC density lipoprotein (HDL). Cleavage of APO-1 from HDL leads to
 CC coagulation of the remaining fraction resulting in opalescence.
 This
 CC observation is a useful tool for qualitatively determining the
 CC concentration of HDL in mammalian sera, low levels of which
 contribute
 CC to coronary heart disease.
 SQ Sequence 1025 AA;

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:20:29 1996; MasPar time 13.87
Seconds

803.752 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-56

Description: (1-441) from US08121713B.pep

Perfect Score: 3231

S e q u e n c e :

MMVLLHAVYSIVFVDVIIK.....FIADIGIGVGMPQMKKILKM 441

1

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3
7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 47.658; Variance 115.370; scale 0.413

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID	Description	
Pred. No.							

1	3231	100.0	441	6	S29921	hypothetical protein	
0.00e+00							

2	2907	90.0	403	6	E42521	A39R protein - vacci	
0.00e+00	3	2137	66.1	295	6	JQ1775	SalL9R protein - vac
0.00e+00	4	1019	31.5	142	6	JQ1776	SalF1R protein - vac
3.01e-146	5	721	22.3	139	6	JQ1846	15R protein - variol
1.39e-95	6	713	22.1	139	6	I36852	A44R protein - vario
3.07e-94	7	608	18.8	122	6	H36852	A43R protein - vario
1.09e-76	8	608	18.8	122	6	JQ1845	14R protein - variol
1.09e-76	9	489	15.1	74	6	G36852	A42R protein - vario
4.24e-57	10	477	14.8	69	6	JQ1844	13R protein - variol
3.80e-55	11	158	4.9	771	11	D49423	semaphorin III precu
8.74e-07	12	157	4.9	772	10	A49069	collapsin - chicken
1.18e-06	13	150	4.6	724	10	C49423	semaphorin II precur
9.21e-06	14	147	4.5	711	13	A49423	semaphorin I precurs
2.20e-05	15	137	4.2	730	10	JH0798	fasciclin IV precurs
3.78e-04	16	125	3.9	656	10	B49423	semaphorin I - fruit
1.01e-02	17	113	3.5	1036	12	S55984	probable homeobox-do
2.31e-01	18	109	3.4	522	11	A31556	glucose transport pr
6.30e-01	19	110	3.4	523	11	S05319	glucose transport pr
4.91e-01	20	110	3.4	523	11	S06920	glucose transport pr
4.91e-01	21	106	3.3	113	11	B44887	glucose transporter
1.31e+00	22	102	3.2	443	9	B26696	hypothetical protein
3.43e+00	23	104	3.2	445	9	E22845	hypothetical protein
2.13e+00	24	101	3.1	169	12	S51910	G4 protein - Saurole
4.34e+00	25	100	3.1	343	3	MMBEA5	cell fusion protein
5.49e+00	26	100	3.1	343	3	A42746	cell fusion protein
5.49e+00	27	101	3.1	499	4	S01528	NADH dehydrogenase (
4.34e+00	28	101	3.1	499	1	DELVN4	NADH dehydrogenase (
4.34e+00							

29	101	3.1	500	6	C34955	hypothetical protein
4.34e+00						
30	99	3.1	609	4	S04757	NADH dehydrogenase (
6.93e+00						
31	101	3.1	2749	2	ACMSIT	inositol-trisphospha
4.34e+00						
32	97	3.0	263	9	S52492	hypothetical protein
1.10e+01						
33	96	3.0	325	2	TVHUAS	transforming protein
1.38e+01						
34	96	3.0	365	1	S26019	ubiquinol--cytochrom
1.38e+01						
35	96	3.0	406	12	S34396	type I site-specific
1.38e+01						
36	96	3.0	443	4	F30010	NADH dehydrogenase (
1.38e+01						
37	96	3.0	474	8	JX0301	ubiquinol--cytochrom
1.38e+01						
38	96	3.0	587	10	S41409	envelysin (EC 3.4.24
1.38e+01						
39	96	3.0	587	10	S12805	envelysin (EC 3.4.24
1.38e+01						
40	97	3.0	656	9	S49745	hypothetical protein
1.10e+01						
41	96	3.0	712	3	BVBRCB	cyaB protein - Borde
1.38e+01						
42	97	3.0	1256	3	A43829	muramidase-released
1.10e+01						
43	96	3.0	2670	6	A46719	inositol 1,4,5-trisp
1.38e+01						
44	97	3.0	2734	6	B36579	inositol 1,4,5-triph
1.10e+01						
45	97	3.0	2749	6	A36579	inositol 1,4,5-triph
1.10e+01						

ALIGNMENTS

```

RESULT      1
ENTRY       S29921      #type complete
TITLE       hypothetical protein 15 - vaccinia virus
ORGANISM    #formal_name vaccinia virus
DATE        20-Feb-1995 #sequence_revision 20-Feb-1995
#text_change
            20-Feb-1995
ACCESSIONS  S29921
REFERENCE   S29907
#authors    Amegadzie, B.Y.
#submission submitted to the EMBL Data Library, January 1991
#accession  S29921
##status    preliminary
##molecule_type DNA
##residues  1-441 ##label AME
##cross-references EMBL:X57318

```

SUMMARY #length 441 #molecular-weight 50185 #checksum 6034

Query Match 100.0%; Score 3231; DB 6; Length 441;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 441; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

D b 1
mmvllhavysivfvdviiikvqryindiltldifylfkmiplllfilfyfangiewhkfet 60

||||| 1
Q y 1
MMVLLHAVYSIVFVDVIIIKVQRYINDILTLDIFYLFKMIPLLLFILFYFANGIEWHKFET 60

D b 6 1
seeiistyllddvlytgvngavytfsnnklnktgltnnnnyittsikvedadkdtlvcgtn 120

||||| 1
Q y 6 1
SEEIISTYLLDDVLYTGAVYTFSSNNKLNKTGLTNNNYITTSIKVEDADKDTLVCGTN 120

D b 1 2 1
ngnpkcwkidgsddpkhrgrgyapyqnskvttiishngcvlsdiniskegikrwrfdgpc 180

||||| 1
Q y 1 2 1
NGNPKCWKIDGSDDPKHRGRGYAPYQNSKVTTIISHNGCVLS DINISKEGIKRWRFRDGPC 180

D b 1 8 1
gydlytadnvipkdglrgafvdkdgt ydkvyilftdtigskrivkipyiaqmclndeggp 240

||||| 1
Q y 1 8 1
GYDLYTADNVIPKDGLRGAFVDKDGTYDKVYILFTDTIGSKRIVKIPYIAQMCLNDEGGP 240

D b 2 4 1
sslsshrwstflkvelecdidgrsyqrqiihsrtiktdndtilyvffdspysksalctysm 300

||||| 1
Q y 2 4 1
SSLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIKTDNDTILYVFFDSPYSKSALCTYSM 300

D b 3 0 1
ntikqsfstsklegytqqlpspasgiclpagkvvphttfeviekynvlddiikplsnqpi 360

||||| 1
Q y 3 0 1
NTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTT FEVIEKYNVLDDI IKPLSNQPI 360

D b 3 6 1
fegpsgvkwfdikekenehreyriyfikensiysfdtkskqtrssqvdarlfsvmvtskp 420

|||||

Q y 3 6 1
FEGPSGVKWFDIKEKENEHREYRIYFIKENSISYFDTKSKQTRSSQVDARLFSVMVTSKP 420

Db 421 lfiadigigvgmpqmkkilkm 441
||||||||||||||||
Qy 421 LFIADIGIGVGMPQMKKILKM 441

RESULT 2
ENTRY E42521 #type complete
TITLE A39R protein - vaccinia virus (strain Copenhagen)
ORGANISM #formal_name vaccinia virus
#note host Homo sapiens (man)
DATE 09-Nov-1990 #sequence_revision 09-Nov-1990
#text_change 08-Apr-1994
ACCESSIONS E42521
REFERENCE A33172
#authors Johnson, G.P.
#submission submitted to GenBank, June 1990
#accession E42521
##status preliminary
##molecule_type DNA
##residues 1-403 ##label JOH
SUMMARY #length 403 #molecular-weight 45741 #checksum
8167

Query Match 90.0%; Score 2907; DB 6; Length 403;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 399; Conservative 2; Mismatches 2; Indels 0;
Gaps 0;

D b 1
mipllfilfyfangiewhkfetseeiistyllddvlytgvngavytfsnnklnktgltnn 60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Q y 3 9
MIPLLFILFYFANGIEWHKFETSEEIISTYLLDDVLYTGVNGAVYTFSNNKLNKTGLTNN 98

D b 6 1
nyittsikvedadkdtlvcgtnnngnpkcwkidgsddpkhrgrgyapyqnskvttiisynec 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||

| |
Q y 9 9
NYITTSIKVEDADKDTLVCGTNNGNPKCWKIDGSDDPKHRGRGYAPYQNSKVTTIISHNGC 158

D b 1 2 1
vlsdiniskegikrwrfrdgpcgydlytadnvipkdglrgafvdkdgtvyilftdti 180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Q y 1 5 9
VLSDINISKEGIKRWRRFDGPCGYDLYTADNVIPKDGLRGAFVDKGTVDKVIILFTDTI 218

D b 1 8 1

```

gskrivkipyiaqmclndeggpsslsshrwstflkvelecdidgrsyqrqiihsrtiktdn 240
|||||
Q          y          2          1          9
GSKRIVKIPYIAQMCLNDEGGPSSLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIKTDN 278

D          b          2          4          1
dtilyvffdspysksalctysmntikqsfstsklegytqqlpspagiclpagkvvshtt 300
|||||
Q          y          2          7          9
DTILYVFFDSPYSKSALCTYSMNTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTT 338

D          b          3          0          1
feviekynvlddiikplsnqpifegpsgvkwfdikekenehreyriyifikensiysfjdk 360
|||||
Q          y          3          3          9
FEVIEKYNVLDDIIKPLSNQPIFEGPSGVKWFDIKEKENEHREYRIYFIKENSIIYSFDTK 398

Db      361 skqtrssqvdarlfsvmvtskplfiadigigvgmpqmkkilkm 403
      |||||
Qy      399 SKQTRSSQVDARLFSVMVTSKPLFIADIGIGVGMPQMKKILKM 441

```

```

RESULT      3
ENTRY      JQ1775      #type complete
TITLE      Sall9R protein - vaccinia virus (strain WR)
ORGANISM    #formal_name vaccinia virus
DATE        30-Sep-1993 #sequence_revision 30-Sep-1993
#text_change
          30-Sep-1993
ACCESSIONS  JQ1775
REFERENCE    JQ1767
      #authors    Smith, G.L.; Chan, Y.S.; Howard, S.T.
      #journal    J. Gen. Virol. (1991) 72:1349-1376
      #title      Nucleotide sequence of 42kbp of vaccinia virus
strain WR from
          near the right inverted terminal repeat.
      #accession  JQ1775
      ##molecule_type DNA
      ##residues  1-295 ##label SMI
      ##cross-references DDBJ:D11079
SUMMARY      #length 295 #molecular-weight 33699 #checksum
7539

```

```

Query Match      66.1%; Score 2137; DB 6; Length 295;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 290; Conservative 0; Mismatches 0; Indels 1;
Gaps 1;

```

```

D          b          1
mmvllhavysivfdviiikvqryindiltldifylfkmiplllfilfyfangiewhkfet 60

```

```

|||||
Q      Y      1
MMVLLHAVYSIVFVDVIIIKVQRYINDILTLDIFYLFKMIPLLFIIFYFANGIEWHKFET 60 1

D      b      6      1
seeiistyllddvlytgvngavytfsnnklnktgltnnnyittsikvedadkdtlvcgtn 120 1

|||||
Q      Y      6      1
SEEIISTYLLDDVLYTGVNGAVYTFSNNKLNKTGLTNNNYITTSIKVEDADKDTLVCGTN 120 1

D      b      1      2      1
ngnpkckidgsddpkhrgrgyapyqnskvttiishngcvlsdiniskegikrwrfrdgpc 180 1

|||||
Q      Y      1      2      1
NGNPKCWKIDGSDDPKHRGRGYAPYQNSKVTIISHNGCVLS DINISKEGIKRWRFRDGPC 180 1

D      b      1      8      1
gydlytadnvipkdglrgafvdkdgydkvyilftdtigskrivkipyiaqmclndeggp 240 1

|||||
Q      Y      1      8      1
GYDLYTADNVIPKDGLRGAFVDKDGTYDKVYILFTDTIGSKRIVKIPYIAQMCLNDEGGP 240 1

Db      241  sslsshrwstflkvelecdidgrsyqrqihhsrtiktdndtilyvffdsphy 291
Qy      241  SSLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIKTDNDTILYVFFDSP-Y 290

```

```

RESULT      4
ENTRY      JQ1776      #type complete
TITLE      SalFlR protein - vaccinia virus (strain WR)
ORGANISM    #formal name vaccinia virus
DATE        30-Sep-1993 #sequence_revision 30-Sep-1993
#text_change
          30-Sep-1993
ACCESSIONS  JQ1776
REFERENCE    JQ1767
#authors     Smith, G.L.; Chan, Y.S.; Howard, S.T.
#journal      J. Gen. Virol. (1991) 72:1349-1376
#title        Nucleotide sequence of 42kbp of vaccinia virus
strain WR from
          near the right inverted terminal repeat.
#accession    JQ1776
##molecule_type DNA
##residues    1-142 ##label SMI
##cross-references DDBJ:D11079
SUMMARY      #length 142 #molecular-weight 16133 #checksum
5310

```

```

Query Match      31.5%; Score 1019; DB 6; Length 142;
Best Local Similarity 99.3%; Pred. No. 3.01e-146;
Matches 141; Conservative 1; Mismatches 0; Indels 0;

```


Gaps 0;

D b 1
mntikqsfstsklegytqqlpspagiclpagkvvphttfeviekynvlddiikplsnqp 60

|||||:|||||
Q y 3 0 0
MNTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTTFFEVIEKYNVLDDIIKPLSNQP 359

D b 6 1
ifegpsgvkwfdikekenehreyriyifikensiysfdtkskqtrssqvdarlfsvmvtsk 120

|||||:|||||
Q y 3 6 0
IFEGPSGVKWFEDIKEKENEHREYRIYFIKENSIYSFDTKSKQTRSSQVDARLFSVMVTSK 419

Db 121 plfiadigigvgmpqmkilkm 142

Qy 420 PLFIADIGIGVGMPQMKILKM 441

RESULT 5

ENTRY JQ1846 #type complete
TITLE 15R protein - variola major virus
ORGANISM #formal_name variola major virus
DATE 03-May-1994 #sequence_revision 03-May-1994
#text_change

03-May-1994

ACCESSIONS JQ1846

REFERENCE JQ1832

#authors Aguado, B.; Selmes, I.P.; Smith, G.L.

#journal J. Gen. Virol. (1992) 73:2887-2902

#title Nucleotide sequence of 21.8 kbp of variola major virus strain

Harvey and comparison with vaccinia virus.

#accession JQ1846

##molecule_type DNA

##residues 1-139 ##label AGU

##experimental_source strain Harver

SUMMARY #length 139 #molecular-weight 16349 #checksum 2421

Query Match 22.3%; Score 721; DB 6; Length 139;
Best Local Similarity 75.4%; Pred. No. 1.39e-95;
Matches 104; Conservative 9; Mismatches 25; Indels 0;
Gaps 0;

D b 1
mntikqsfstsnwediqsnyclqllyvyqlekvvphtfdvieqynvldniikplsnqp 60

|||||:|||||
Q y 3 0 0
MNTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTTFFEVIEKYNVLDDIIKPLSNQP 359

D ifkgpsdvkwfdikekenehrkyriyfikentiysfntkskqtrssqvdaqlfsvmvtsk 120
 |||||:||||:|||||:|||||:|||||
 Q IFEGPSGVKWFDIKEKENEHREYRIYFIKENSIYSFDTKSKQTRSSQVDARLFSVMVTSK 419

```

RESULT      6
ENTRY       I36852      #type complete
TITLE       A44R protein - variola virus (strain India-1967)
ORGANISM    #formal_name variola virus
DATE        30-Sep-1993 #sequence_revision 30-Sep-1993
#text_change
            30-Sep-1993
ACCESSIONS  I36852
REFERENCE   A36859
#authors    Blinov, V.M.
#submission submitted to GenBank, November 1992
#description not shown.
#accession  I36852
##status    preliminary
##molecule_type DNA
##residues  1-139 ##label BLI
##cross-references GB:X69198
SUMMARY     #length 139  #molecular-weight 16409  #checksum
1680

```

```

D                                     b                                     1
mntikqsfstsnwedigsnyclqllvyvyqlekvvphntfdviegynvldniikplfnqp 60
      ||||| |||||: |                                     ||||| ||: ||| |||||: |||||
|||
Q                                     y                                     3                                     0                                     0
MNTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTTFEVIEKYNVLDDIIKPLSNQP 359

```

```

Db      121 plfiadigiegvgmprikn 138
        ||||| | | | | : : :
Qy      420 PLFIADIGIGVGMPQMKK 437

```

RESULT 7
 ENTRY H36852 #type complete
 TITLE A43R protein - variola virus (strain India-1967)
 ORGANISM #formal_name variola virus
 DATE 30-Sep-1993 #sequence_revision 30-Sep-1993
 #text_change

30-Sep-1993

ACCESSIONS H36852
 REFERENCE A36859
 #authors Blinov, V.M.
 #submission submitted to GenBank, November 1992
 #description not shown.
 #accession H36852
 ##status preliminary
 ##molecule_type DNA
 ##residues 1-122 ##label BLI
 ##cross-references GB:X69198

SUMMARY #length 122 #molecular-weight 14301 #checksum 8838

Query Match 18.8%; Score 608; DB 6; Length 122;
 Best Local Similarity 91.5%; Pred. No. 1.09e-76;
 Matches 86; Conservative 3; Mismatches 4; Indels 1;
 Gaps 1;

D b 3
 y-lytadnvipkdglqgafvdkdgytdkvylftvtigtiskrivkipyiaqmclndecgps 61
 |||||:|||||
 |||
 Q y 1 8 2
 YDLYTADNVIPKDGRLRGAFVDKDGTYDKVYILFTDTIGSKRIVKIPYIAQMCLNDEGGPS 241

Db 62 slsshrwstllkvelecdidgrsysqinhsktik 95
 |||||:|||||
 Qy 242 SLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIK 275

RESULT 15
 ENTRY JH0798 #type complete
 TITLE fasciclin IV precursor - American bird grasshopper
 ORGANISM #formal_name Schistocerca americana #common_name American
 bird grasshopper
 DATE 30-Sep-1993 #sequence_revision 30-Sep-1993
 #text_change
 11-Apr-1995
 ACCESSIONS JH0798
 REFERENCE JH0798
 #authors Kolodkin, A.; Matthes, D.J.; O'Connor, T.P.;
 Patel, N.H.;
 Admon, A.; Bentley, D.; Goodman, C.S.
 #journal Neuron (1992) 9:831-845
 #title Fasciclin IV: Sequence, expression, and function
 during

```

          growth cone guidance in the grasshopper embryo.
#accession      JH0798
  ##molecule_type mRNA
  ##residues      1-730 ##label KOL
  ##cross-references GB:L00709
  ##experimental_source embryo
COMMENT      This protein plays a role in growth cone guidance in the
developing
              central nervous system.
KEYWORDS      glycoprotein; transmembrane protein
FEATURE
  1-22          #domain signal sequence #status predicted
#label SIG\
  23-730        #product fasciclin IV #status predicted
#label MAT\
  23-627        #domain extracellular #label EXT\
  628-652      #domain transmembrane #label TMM\
  653-730      #domain intracellular #label INT\
  44,71,163,267,360,
  539          #binding_site carbohydrate (Asn) (covalent)
#status
              predicted
SUMMARY      #length 730 #molecular-weight 81214 #checksum
5881

```

```

Query Match      4.2%; Score 137; DB 10; Length 730;
Best Local Similarity 31.1%; Pred. No. 3.78e-04;
Matches      38; Conservative 29; Mismatches 40; Indels 15;
Gaps 11;

```

```

D          b          2          0          8
fvntmeyndfiffiffretaveyincgkaiysrvarvckhdkggphqfgd-rwtsflksrl 266
      ||:      | ::::| :| :      | | :|::| :| ||| :: ||::|||
|
Q          y          2          0          0
FVDKDGTYDKVYILFTDT-IGSKRIVKIPY--IAQMCLNDEGGPSSLSSHRWSTFLKVEL 256

D          b          2          6          7
ncsvpgdypfyfneiqtstsiiegnyggqvekliygvtfttpvnsiggsavcafsmksile 326
      :| : | : | :| : | :      : | | | :| |
||:|::||::| :
Q          y          2          5          7
ECDIDGR-S-YRQIIHSRT--IKTD-NDTI--L-YVFFDSP-YSK--SALCTYSMNTIKQ 305

Db      327 sf 328
      ||
Qy      306 SF 307

```

```

Search completed: Thu May 16 15:21:52 1996
Job time : 83 secs.

```

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:19:36 1996; MasPar time 8.33
Seconds

811.743 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-56

Description: (1-441) from US08121713B.pep

Perfect Score: 3231

S e q u e n c e :

MMVLLHAVYSIVFVDVIIK.....FIADIGIGVGMPQMKKILKM 441

1

Scoring table: PAM 150
Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot31
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8

Statistics: Mean 49.812; Variance 96.211; scale 0.518

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
Pred. No.							
1	3231	100.0	441	7	VA39_VACCV	PROTEIN A39.	
0.00e+00							
2	2907	90.0	403	7	VA39_VACCC	PROTEIN A39.	

0.00e+00	3	109	3.4	522	3	GTR2_RAT	GLUCOSE TRANSPORTER T
7.87e-02	4	110	3.4	523	3	GTR2_MOUSE	GLUCOSE TRANSPORTER T
5.84e-02	5	100	3.1	343	2	CELF_HSVEB	CELL FUSION PROTEIN P
1.05e+00	6	100	3.1	365	2	CYB_ASCSU	CYTOCHROME B (EC 1.10
1.05e+00	7	101	3.1	499	5	NU4C_MARPO	NADH-PLASTOQUINONE OX
7.93e-01	8	101	3.1	500	5	PCD6_MOUSE	PROTEIN PCD-6 (FRAGME
7.93e-01	9	99	3.1	609	5	NU5M_RAT	NADH-UBIQUINONE OXIDO
1.38e+00	10	101	3.1	2749	4	IP3R_MOUSE	INOSITOL 1,4,5-TRISPH
7.93e-01	11	96	3.0	325	4	MAS_HUMAN	MAS PROTO-ONCOGENE.
3.13e+00	12	96	3.0	587	3	HE_PARLI	HATCHING ENZYME PRECU
3.13e+00	13	96	3.0	712	2	CYAB_BORPE	CYCLOLYSIN SECRETION
3.13e+00	14	97	3.0	1256	5	MRP_STRSU	MURAMIDASE-RELEASED P
2.39e+00	15	97	3.0	2749	4	IP3R_RAT	INOSITOL 1,4,5-TRISPH
2.39e+00	16	94	2.9	91	8	YVBG_VACCC	HYPOTHETICAL 11.2 KD
5.32e+00	17	94	2.9	198	2	COP_CLOPE	COPY NUMBER PROTEIN (
5.32e+00	18	95	2.9	338	1	ANX2_HUMAN	ANNEXIN II (LIPOCORTI
4.09e+00	19	94	2.9	447	5	NU4M_APIME	NADH-UBIQUINONE OXIDO
5.32e+00	20	94	2.9	499	3	GAB_LYMST	GAMMA-AMINOBUTYRIC-AC
5.32e+00	21	93	2.9	506	2	CP5I_CANTR	CYTOCHROME P450 LIIB1
6.92e+00	22	94	2.9	532	8	YHB7_YEAST	HYPOTHETICAL 61.2 KD
5.32e+00	23	94	2.9	707	3	HLYB_ACTPL	HAEMOLYSIN SECRETION
5.32e+00	24	94	2.9	1951	2	CIN3_RAT	SODIUM CHANNEL PROTEI
5.32e+00	25	91	2.8	167	8	YPA2_ASCIM	HYPOTHETICAL 19.7 KD
1.16e+01	26	90	2.8	204	1	BACR_HALHM	BACTERIORHODOPSIN (BR
1.50e+01	27	92	2.8	259	1	BAC2_HALS2	ARCHAERHODOPSIN 2 PRE
8.97e+00	28	90	2.8	267	1	BLO5_PSEAE	BETA-LACTAMASE OXA-5
1.50e+01	29	92	2.8	277	1	BLO7_ECOLI	BETA-LACTAMASE OXA-7

8.97e+00	30	92	2.8	291	7	SUAR_RAT	ARYL SULFOTRANSFERASE
8.97e+00	31	90	2.8	303	8	YE15_YEAST	HYPOTHETICAL 33.7 KD
1.50e+01	32	91	2.8	413	3	HEMO_HYACE	HEMOLIN PRECURSOR (P4
1.16e+01	33	90	2.8	483	8	XYLB_STAXY	XYLULOSE KINASE (EC 2
1.50e+01	34	91	2.8	501	5	PHR1_SINAL	DEOXYRIBODIPYRIMIDINE
1.16e+01	35	92	2.8	554	5	NU5M_APIME	NADH-UBIQUINONE OXIDO
8.97e+00	36	91	2.8	580	5	P69_MYCHR	TRANSPORT SYSTEM PERM
1.16e+01	37	91	2.8	702	1	ATI1_VARV	81 KD A-TYPE INCLUSIO
1.16e+01	38	90	2.8	707	3	HLYB_ECOLI	HAEMOLYSIN SECRETION
1.50e+01	39	90	2.8	707	3	HLY2_ECOLI	HAEMOLYSIN SECRETION
1.50e+01	40	91	2.8	718	5	NAMI_CANFA	SODIUM/MYO-INOSITOL C
1.16e+01	41	91	2.8	774	8	VP4_ROTHT	OUTER CAPSID PROTEIN
1.16e+01	42	92	2.8	775	8	VP4_ROTHM	OUTER CAPSID PROTEIN
8.97e+00	43	92	2.8	776	8	VP4_ROTHN	OUTER CAPSID PROTEIN
8.97e+00	44	92	2.8	1522	2	CIN1_LOLBL	SODIUM CHANNEL PROTEI
8.97e+00	45	92	2.8	1590	3	GCN2_YEAST	PROTEIN KINASE GCN2 (
8.97e+00							

ALIGNMENTS

RESULT 1
 ID VA39 VACCV STANDARD; PRT; 441 AA.
 AC P24764;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE PROTEIN A39.
 GN A39R OR (SALL9R AND SALL1R).
 OS VACCINIA VIRUS (STRAIN WR).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE;
 CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91310644
 RA AMEGADZIE B.Y., AHN B.-Y., MOSS B.;
 RL J. BIOL. CHEM. 266:13712-13718(1991).

→ Deleted sequence not
 in this article

RN [2]
 RP SEQUENCE FROM N.A.
 RM 91259063
 RA SMITH G.L., CHAN Y.S., HOWARD S.T.;
 RL J. GEN. VIROL. 72:1349-1376(1991).
 CC -!- CAUTION: IN REF.2 THIS SEQUENCE IS IN TWO PART (SALL9R AND
 CC SALF1R) DUE TO THE PRESENCE OF FRAMESHIFT COMPARED TO THAT
 OF
 CC REF.1.
 DR EMBL; M61187; VACRNAPS.
 DR EMBL; D11079; PXVACRHF.
 DR EMBL; X57318; VVDNA.
 DR PIR; S29921; S29921.
 SQ SEQUENCE 441 AA; 50185 MW; 1082834 CN;

Query Match 100.0%; Score 3231; DB 7; Length 441;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 441; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

D b 1
 mmvllhavysivfvdviiiikvqryindiltldifylfkmiplllfilfyfangiewhkfet 60

|
 Q y 1
 MMVLLHAVYSIVFVDVIIIKVQRYINDILTLDFYLFKMIPLLLFILFYFANGIEWHKKFET 60

D b 6 1
 seeiistyllddvlytgvngavytfsnnklnktgltnnnnyittsikvedadkdtlvcgtn 120

|
 Q y 6 1
 SEEIISTYLLDDVLYTGAVYTFSSNNKLNKTGLTNNNYITTSIKVEDADKDTLVCGTN 120

D b 1 2 1
 ngnpkcwkidgsddpkhrgrgyapyqnskvttiishngcvlsdiniskegikrwrfdgpc 180

|
 Q y 1 2 1
 NGNPKCWKIDGSDDPKHRGRGYAPYQNSKVTIISHNGCVLSDINISKEGIKRWRFDGPC 180

D b 1 8 1
 gydlytadnvipkdglrgafvdkdgtvyilftdtigskrivkipyiaqmclndeggp 240

|
 Q y 1 8 1
 GYDLYTADNVIPKDGLRGAFVDKDGTYDKVYILFTDTIGSKRIVKIPYIAQMCLNDEGGP 240

D b 2 4 1
 sslsshrwstflkvelecdidgrsyqrqiihsrtiktdndtilyvffdspysksalctysm 300

|
 Q y 2 4 1
 SSLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIKTDNDTILYVFFDSPYSKSALCTYSM 300


```

miplllfilfyfangiewhkfetseeiistyllddvlytgvngavytfsnnklnktgltnn 60
|||||
Q      y      3      9
MIPLLFILFYFANGIEWHKFETSEEIISTYLLDDVLYTGVNGAVYTFSSNNKLNKTGLTNN 98

D      b      6      1
nyittsikvedadkdtlvcgtnnngnpkcwkidgsddpkhrgrgyapyqnskvttiisynec 120
|||||
| |
Q      y      9      9
NYITTSIKVEDADKDTLVCGTNNGNPKCWKIDGSDDPKHRGRGYAPYQNSKVTIISHNGC 158

D      b      1      2      1
vlsdiniskegikrwr rfdgpcgydlytadnvipkdglrgafvdkdgydkvyilftdti 180
|||||
Q      y      1      5      9
VLSDINISKEGIKRWR RFDGPCGYDLYTADNVIPKDGLRGAFVDKDGTYDKVYILFTDTI 218

D      b      1      8      1
gskrivkipyiaqmclndeggpsslsshrwstflkvelecdidgrsyqrqiihsrtiktdn 240
|||||
Q      y      2      1      9
GSKRIVKIPYIAQMCLNDEGGPSSLSSHRWSTFLKVELECDIDGRSYRQIIHSRTIKTDN 278

D      b      2      4      1
dtilyvffdspysksalctysmntikqsfstsklegytqqlpaspagiclpagkvvshtt 300
|||||
Q      y      2      7      9
DTILYVFFDSPYSKSALCTYSMNTIKQSFSTSKLEGYTKQLPSPASGICLPAGKVVPHTT 338

D      b      3      0      1
feviekynvlddiikplsnqpifegpsgvkwfdikekenehreyriyfikensiysfdtk 360
|||||
Q      y      3      3      9
FEVIEKYNVLDDIIKPLSNQPIFEGPSGVKWFDIKEKENEHREYRIYFIKENSIYSFDTK 398

Db      361 skqtrssqvdarlfsvmvtskplfiadigigvgmpqmkkilkm 403
|||
Qy      399 SKQTRSSQVDARLF SVMVTSKPLFIADIGIGVGMPQMKKILKM 441

```

Search completed: Thu May 16 15:20:10 1996
Job time : 34 secs.

Unit. Release 2.1D John F. Collins, Biocomputing Research
Copyright (c) 1993, 1994, 1995 University of Edinburgh,
U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:25:47 1996; MasPar time 9.88
Seconds

612.229 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-58

Description: (1-730) from US08121713B.pep

Perfect Score: 5233

S e q u e n c e :

MRAALVAVAALLWVALHAAA.....KNANSSAENKPIQKVKKTYI 730

1

Scoring table: PAM 150
Gap 11

Searched: 70887 seqs, 8282111 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8 9:part9 10:part10 11:part11 12:part12
13:part13
14:part14

Statistics: Mean 37.226; Variance 160.823; scale 0.231

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description
Pred. No.						
1	5233	100.0	730	13	R71379	Grasshopper semaphori

0.00e+00	2	2394	45.7	712	13	R71384	Tribolium semaphorin
6.86e-220	3	2060	39.4	650	13	R71382	Drosophila semaphorin
8.26e-187	4	963	18.4	724	13	R71383	Drosophila semaphorin
4.89e-79	5	910	17.4	771	13	R71380	Human semaphorin III
6.77e-74	6	659	12.6	477	13	R74175	Human collapsin.
1.00e-49	7	137	2.6	441	13	R71381	Vaccinia virus semaph
2.37e-02	8	101	1.9	122	13	R71385	Variola major virus s
1.07e+01	9	98	1.9	499	12	R60609	Tobamovirus replicati
1.73e+01	10	101	1.9	572	13	R78520	Partial ALK protein.
1.07e+01	11	98	1.9	793	10	R53758	H. influenzae PAK 120
1.73e+01	12	98	1.9	797	10	R53755	H. influenzae b Eagen
1.73e+01	13	98	1.9	797	10	R53756	H. influenzae b Minn
1.73e+01	14	98	1.9	797	10	R53754	H. influenzae b Ca D1
1.73e+01	15	97	1.9	3647	1	R05041	Filamentous haemagglu
2.02e+01	16	95	1.8	348	4	R27662	C-terminal region of
2.77e+01	17	95	1.8	455	4	R23973	Transmembrane form of
2.77e+01	18	96	1.8	516	1	P81987	Recombinant human lys
2.37e+01	19	92	1.8	558	2	P70430	Human neuroleukin.
4.41e+01	20	94	1.8	792	10	R53757	H. influenzae SB33 D1
3.23e+01	21	95	1.8	826	5	R26042	P. yoelii SSP2 antige
2.77e+01	22	94	1.8	1337	14	R85203	huDEP-1.
3.23e+01	23	88	1.7	441	3	P50309	Endoglucanase plasmid
8.10e+01	24	88	1.7	458	14	R79539	Endoglucanase-I prote
8.10e+01	25	88	1.7	458	14	R77263	T. longibrachiatum en
8.10e+01	26	88	1.7	459	3	P50134	Sequence encoded by e
8.10e+01	27	91	1.7	586	2	R07998	Asparagine synthetase
5.14e+01	28	90	1.7	604	10	R51268	Sequence of murine pr

5.98e+01	29	90	1.7	661	2	R07504	Merozite apical-end-1
5.98e+01	30	87	1.7	705	10	R66597	Human L5/3 tumour sup
9.41e+01	31	87	1.7	705	10	R66598	Human L5/3 tumour sup
9.41e+01	32	87	1.7	711	10	R66602	Human L5/3 tumour sup
9.41e+01	33	87	1.7	713	10	R66603	Encoded by full-lengt
9.41e+01	34	91	1.7	768	5	R27683	Rabbit beta-8 intergi
5.14e+01	35	90	1.7	769	1	P80836	Beta subunit of human
5.98e+01	36	90	1.7	769	1	R07113	Recombinant beta-subu
5.98e+01	37	90	1.7	769	4	R24256	Beta-subunit CD18 of
5.98e+01	38	87	1.7	941	2	R07478	Cellulase.
9.41e+01	39	91	1.7	4987	3	R10834	Rianodin receptor.
5.14e+01	40	85	1.6	274	10	R54837	B. lichen subtilisin
1.27e+02	41	85	1.6	600	8	R44614	Thermostable ligase.
1.27e+02	42	86	1.6	622	1	P91632	Rhoptry membrane anti
1.09e+02	43	86	1.6	768	5	R27684	Human beta-8 intergin
1.09e+02	44	86	1.6	1140	14	R72386	XAP-1, part of the DN
1.09e+02	45	85	1.6	3898	2	R06996	Protein characteristi
1.27e+02							

ALIGNMENTS

Search completed: Thu May 16 15:26:23 1996

Job time : 36 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:24:16 1996; MasPar time 21.66
Seconds

851.723 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-58

Description: (1-730) from US08121713B.pep

Perfect Score: 5233

S e q u e n c e :

MRAALVAVAALLWVALHAAA.....KNANSSAENKPIQKVKKTYI 730

1

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3
7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 50.468; Variance 112.049; scale 0.450

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
No.						
Pred. No.						

1	5220	99.8	730	10	JH0798	fasciclin IV precurs
0.00e+00						
2	2403	45.9	711	13	A49423	semaphorin I precurs
0.00e+00						

3	2086	39.9	656	10	B49423	semaphorin I - fruit
0.00e+00						
4	963	18.4	724	10	C49423	semaphorin II precu
1.54e-147						
5	916	17.5	772	10	A49069	collapsin - chicken
8.14e-139						
6	910	17.4	771	11	D49423	semaphorin III precu
1.05e-137						
7	147	2.8	403	6	E42521	A39R protein - vacci
1.63e-05						
8	137	2.6	441	6	S29921	hypothetical protein
3.30e-04						
9	131	2.5	295	6	JQ1775	SalL9R protein - vac
1.90e-03						
10	109	2.1	329	8	H64115	phenylalanine--tRNA
7.89e-01						
11	109	2.1	535	3	WMBEW1	UL21 protein - human
7.89e-01						
12	108	2.1	551	10	S52287	urbain - silkworm
1.02e+00						
13	105	2.0	330	10	B48835	kinesin-like protein
2.18e+00						
14	103	2.0	422	9	S56966	probable membrane pr
3.58e+00						
15	106	2.0	4969	11	A37113	ryanodine receptor,
1.69e+00						
16	103	2.0	5037	10	A54161	ryanodine-binding pr
3.58e+00						
17	101	1.9	122	6	H36852	A43R protein - vario
5.85e+00						
18	101	1.9	122	6	JQ1845	14R protein - variol
5.85e+00						
19	99	1.9	268	9	S14068	seed protein precurs
9.48e+00						
20	99	1.9	268	9	S05471	embryonic abundant p
9.48e+00						
21	98	1.9	268	9	S04136	embryonic abundant p
1.20e+01						
22	97	1.9	268	8	C64138	methionine aminopept
1.52e+01						
23	97	1.9	287	11	S12738	T-cell alloantigen R
1.52e+01						
24	100	1.9	421	10	S53818	XPMC2 protein - Afri
7.45e+00						
25	97	1.9	502	13	B46570	H+-transporting ATP
1.52e+01						
26	99	1.9	535	6	C24187	hypothetical protein
9.48e+00						
27	98	1.9	535	1	EUHUGC	glucosylceramidase (
1.20e+01						
28	101	1.9	564	7	I64134	D-lactate dehydrogen
5.85e+00						
29	98	1.9	797	7	JC4078	D-15 protective surf
1.20e+01						

30	100	1.9	917	2	ACGAE	glutamate receptor p
7.45e+00						
31	97	1.9	1031	9	A29839	RAD2 protein - yeast
1.52e+01						
32	100	1.9	1295	6	A35886	polymorphic epitheli
7.45e+00						
33	98	1.9	1615	3	WMTM8T	180K protein - tomat
1.20e+01						
34	96	1.8	380	4	S40049	glucose-1-phosphate
1.93e+01						
35	96	1.8	380	12	S36625	glucose-1-phosphate
1.93e+01						
36	96	1.8	427	13	S29812	transcobalamin - hum
1.93e+01						
37	96	1.8	515	4	A32931	glucosylceramidase (
1.93e+01						
38	95	1.8	630	11	A39344	tumor-associated muc
2.43e+01						
39	96	1.8	712	9	S45308	SSK1 protein - yeast
1.93e+01						
40	96	1.8	808	8	F64102	protective surface a
1.93e+01						
41	95	1.8	826	12	S27834	sporozoite surface p
2.43e+01						
42	95	1.8	826	9	A45559	sporozoite surface p
2.43e+01						
43	96	1.8	976	9	S45738	pleiotropic drug res
1.93e+01						
44	96	1.8	976	9	S46661	PDR3 protein - yeast
1.93e+01						
45	95	1.8	1264	6	A35175	episialin A - human
2.43e+01						

ALIGNMENTS

RESULT	1
ENTRY	JH0798 #type complete
TITLE	fasciclin IV precursor - American bird grasshopper
ORGANISM	#formal_name Schistocerca americana #common_name
American	
	bird grasshopper
DATE	30-Sep-1993 #sequence_revision 30-Sep-1993
#text_change	
	11-Apr-1995
ACCESSIONS	JH0798
REFERENCE	JH0798
#authors	Kolodkin, A.; Matthes, D.J.; O'Connor, T.P.;
Patel, N.H.;	
	Admon, A.; Bentley, D.; Goodman, C.S.
#journal	Neuron (1992) 9:831-845
#title	Fasciclin IV: Sequence, expression, and function
during	
	growth cone guidance in the grasshopper embryo.


```

#accession      JH0798
##molecule_type mRNA
##residues      1-730 ##label KOL
##cross-references GB:L00709
##experimental_source embryo
COMMENT      This protein plays a role in growth cone guidance in the
developing
              central nervous system.
KEYWORDS      glycoprotein; transmembrane protein
FEATURE
  1-22                #domain signal sequence #status predicted
#label SIG\
  23-730              #product fasciclin IV #status predicted
#label MAT\
  23-627              #domain extracellular #label EXT\
  628-652            #domain transmembrane #label TMM\
  653-730            #domain intracellular #label INT\
  44,71,163,267,360,
  539                #binding_site carbohydrate (Asn) (covalent)
#status
              predicted
SUMMARY      #length 730 #molecular-weight 81214 #checksum
5881

```

```

Query Match          99.8%; Score 5220; DB 10; Length 730;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 729; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;

```

```

RESULT      2
ENTRY      A49423      #type complete
TITLE      semaphorin I precursor - beetle (Tribolium
confusum)
ORGANISM      #formal_name Tribolium confusum
DATE      12-Dec-1994 #sequence_revision 12-Dec-1994
#text_change
              12-Dec-1994
ACCESSIONS      A49423
REFERENCE      A49423
#authors      Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
#journal      Cell (1993) 75:1389-1399
#title      The Semaphorin genes encode a family of
transmembrane and
              secreted growth cone guidance molecules.
#accession      A49423
##status      preliminary
##residues      1-711 ##label KOL
##cross-references GB:L26080
SUMMARY      #length 711 #molecular-weight 79623 #checksum
4552

```

```

Query Match          45.9%; Score 2403; DB 13; Length 711;
Best Local Similarity 50.6%; Pred. No. 0.00e+00;
Matches 361; Conservative 145; Mismatches 178; Indels 29;

```

Gaps 23;

```
RESULT      3
ENTRY       B49423      #type fragment
TITLE       semaphorin I - fruit fly (Drosophila melanogaster)
(fragment)
ORGANISM     #formal_name Drosophila melanogaster
DATE         06-Jan-1995 #sequence_revision 06-Jan-1995
#text_change
              27-Jan-1995
ACCESSIONS   B49423
REFERENCE     A49423
#authors     Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
#journal      Cell (1993) 75:1389-1399
#title        The Semaphorin genes encode a family of
transmembrane and
              secreted growth cone guidance molecules.
#accession    B49423
##status      preliminary; not compared with conceptual
translation
##molecule_type mRNA
##residues    1-656 ##label KOL
##cross-references GB:L26082
##note        nucleotide sequence not given
GENETICS
#gene         semaI
SUMMARY       #length 656 #checksum 3860

Query Match          39.9%; Score 2086; DB 10; Length 656;
Best Local Similarity 52.4%; Pred. No. 0.00e+00;
Matches 300; Conservative 134; Mismatches 111; Indels 28;
Gaps 17;
```

```
RESULT      4
ENTRY       C49423      #type complete
TITLE       semaphorin II precursor - fruit fly (Drosophila
melanogaster)
ORGANISM     #formal_name Drosophila melanogaster
DATE         06-Jan-1995 #sequence_revision 06-Jan-1995
#text_change
              27-Jan-1995
ACCESSIONS   C49423
REFERENCE     A49423
#authors     Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
#journal      Cell (1993) 75:1389-1399
#title        The Semaphorin genes encode a family of
transmembrane and
              secreted growth cone guidance molecules.
#accession    C49423
##status      preliminary; not compared with conceptual
translation
##molecule_type mRNA
##residues    1-724 ##label KOL
##cross-references GB:L26083.
```

##note nucleotide sequence not given
GENETICS
#gene sema II
SUMMARY #length 724 #molecular-weight 82971 #checksum
6082

Query Match 18.4%; Score 963; DB 10; Length 724;
Best Local Similarity 37.5%; Pred. No. 1.54e-147;
Matches 199; Conservative 119; Mismatches 159; Indels 53;
Gaps 32;

RESULT 5
ENTRY A49069 #type complete
TITLE collapsin - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 07-Apr-1994 #sequence_revision 07-Apr-1994
#text_change
07-Apr-1994
ACCESSIONS A49069
REFERENCE A49069
#authors Luo, Y.; Raible, D.; Raper, J.A.
#journal Cell (1993) 75:217-227
#title Collapsin: a protein in brain that induces the
collapse and
paralysis of neuronal growth cones.
#accession A49069
##status preliminary; not compared with conceptual
translation
##molecule_type mRNA
##residues 1-772 ##label LUO
##cross-references GB:U02528
SUMMARY #length 772 #molecular-weight 88867 #checksum
9712

Query Match 17.5%; Score 916; DB 10; Length 772;
Best Local Similarity 35.1%; Pred. No. 8.14e-139;
Matches 176; Conservative 124; Mismatches 160; Indels 42;
Gaps 30;

RESULT 6
ENTRY D49423 #type complete
TITLE semaphorin III precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995
#text_change
27-Jan-1995
ACCESSIONS D49423
REFERENCE A49423
#authors Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
#journal Cell (1993) 75:1389-1399
#title The Semaphorin genes encode a family of
transmembrane and
secreted growth cone guidance molecules.
#accession D49423

```
##status      preliminary
##molecule_type mRNA
##residues    1-771 ##label KOL
##cross-references GB:L26081
##note        nucleotide sequence not given
SUMMARY      #length 771 #molecular-weight 88889 #checksum
6249
```

```
Query Match      17.4%; Score 910; DB 11; Length 771;
Best Local Similarity 34.8%; Pred. No. 1.05e-137;
Matches 174; Conservative 120; Mismatches 168; Indels 38;
Gaps 27;
Search completed: Thu May 16 15:25:30 1996
Job time : 74 secs.
```

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:23:09 1996; MasPar time 12.67
Seconds

883.721 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-58

Description: (1-730) from US08121713B.pep

Perfect Score: 5233

S e q u e n c e :

MRAALVAVALLWVALHAAA.....KNANSSAENKPIQKVKKTYI 730

1

Scoring table: PAM 150
Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot31
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8

Statistics: Mean 52.570; Variance 88.484; scale 0.594

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Match	Query Length	DB	ID	Description
No.						
Pred. No.						
1	147	2.8	403	7	VA39_VACCC	PROTEIN A39.
6.42e-08						
2	137	2.6	441	7	VA39_VACCV	PROTEIN A39.

3.01e-06 3	109	2.1	535	7	UL21_HSV11	PROTEIN UL21.
5.86e-02 4	106	2.0	4969	6	RYNC_RABIT	RYANODINE RECEPTOR, C
1.53e-01 5	99	1.9	268	2	EA30_VICFA	EMBRYONIC ABUNDANT PR
1.32e+00 6	98	1.9	268	2	EA92_VICFA	EMBRYONIC ABUNDANT PR
1.78e+00 7	97	1.9	287	6	RT6H_MOUSE	T-CELL DIFFERENTIATIO
2.38e+00 8	97	1.9	502	1	ATPB_CYTLY	ATP SYNTHASE BETA CHA
2.38e+00 9	99	1.9	535	7	UL21_HSV1E	PROTEIN UL21.
1.32e+00 10	100	1.9	917	3	GLRK_LYMST	GLUTAMATE RECEPTOR PR
9.79e-01 11	97	1.9	1031	6	RAD2_YEAST	DNA REPAIR PROTEIN RA
2.38e+00 12	98	1.9	1615	6	RRPO_TMVTO	PUTATIVE RNA-DIRECTED
1.78e+00 13	92	1.8	252	3	ETFB_PARDE	ELECTRON TRANSFER FLA
9.91e+00 14	94	1.8	256	4	LP1_BOMMO	LOW MOLECULAR LIOPRO
5.66e+00 15	94	1.8	264	6	POTC_ECOLI	SPERMIDINE/PUTRESCINE
5.66e+00 16	93	1.8	273	4	KSGA_ECOLI	METHYLTRANSFERASE (ME
7.50e+00 17	95	1.8	287	7	UL24_ILTVT	PROTEIN UL24 HOMOLOG
4.26e+00 18	92	1.8	289	6	ROB_ECOLI	RIGHT ORIGIN-BINDING
9.91e+00 19	93	1.8	296	6	RFBF_SHIFL	DTDP-RHAMNOSYL TRANSF
7.50e+00 20	94	1.8	318	8	YAD5_CLOAB	HYPOTHETICAL 36.9 KD
5.66e+00 21	93	1.8	327	7	SYFA_ECOLI	PHENYLALANYL-TRNA SYN
7.50e+00 22	94	1.8	329	1	ARAF_ECOLI	L-ARABINOSE-BINDING P
5.66e+00 23	96	1.8	380	3	GLGC_BACSU	GLUCOSE-1-PHOSPHATE A
3.19e+00 24	92	1.8	398	8	YKU0_YEAST	HYPOTHETICAL 46.9 KD
9.91e+00 25	93	1.8	427	7	TCO2_HUMAN	TRANSCOBALAMIN II PRE
7.50e+00 26	93	1.8	448	3	GNTP_BACSU	GLUCONATE PERMEASE.
7.50e+00 27	96	1.8	515	3	GLCM_MOUSE	GLUCOSYLCERAMIDASE PR
3.19e+00 28	93	1.8	536	3	GLCM_HUMAN	GLUCOSYLCERAMIDASE PR
7.50e+00 29	92	1.8	558	3	G6PI_HUMAN	GLUCOSE-6-PHOSPHATE I

9.91e+00	30	95	1.8	630	5	MUC1_MOUSE	MUCIN 1 PRECURSOR (PO
4.26e+00	31	93	1.8	659	8	YHH5_YEAST	HYPOTHETICAL 75.9 KD
7.50e+00	32	94	1.8	661	8	YK65_CAEEL	HYPOTHETICAL 73.3 KD
5.66e+00	33	92	1.8	750	7	SYMC_YEAST	METHIONYL-TRNA SYNTH
9.91e+00	34	92	1.8	769	4	ITB2_BOVIN	CELL SURFACE ADHESION
9.91e+00	35	95	1.8	826	7	SSP2_PLAYO	SPOROZOITE SURFACE PR
4.26e+00	36	96	1.8	976	5	PDR3_YEAST	PLEIOTROPIC DRUG RESI
3.19e+00	37	95	1.8	1255	5	MUC1_HUMAN	MUCIN 1 PRECURSOR (PO
4.26e+00	38	92	1.8	3898	6	POLG_BVDVS	GENOME POLYPROTEIN.
9.91e+00	39	91	1.7	193	7	VA57_VACCC	GUANYLATE KINASE HOMO
1.31e+01	40	91	1.7	337	3	G3PX_HORVU	GLYCERALDEHYDE 3-PHOS
1.31e+01	41	91	1.7	337	3	G3PC_MAIZE	GLYCERALDEHYDE 3-PHOS
1.31e+01	42	91	1.7	1286	1	AIDA_ECOLI	ADHESIN AIDA-I PRECUR
1.31e+01	43	91	1.7	1579	5	PEP1_YEAST	VACUOLAR PROTEIN SORT
1.31e+01	44	91	1.7	1615	6	RRPO_TMVKR	PUTATIVE RNA-DIRECTED
1.31e+01	45	91	1.7	2473	7	TOR2_YEAST	PHOSPHATIDYLINOSITOL
1.31e+01							

ALIGNMENTS

RESULT 1
 ID VA39 VACCC STANDARD; PRT; 403 AA.
 AC P21062;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE PROTEIN A39.
 GN A39R.
 OS VACCINIA VIRUS (STRAIN COPENHAGEN).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE;
 CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91021027
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW
 J.P.,

RA PAOLETTI E.;
 RL VIROLOGY 179:247-266(1990).
 RN [2]
 RP COMPLETE GENOME.
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
 RA PAOLETTI E.;
 RL VIROLOGY 179:517-563(1990).
 DR EMBL; M35027; PXVACCG.
 DR PIR; E42521; E42521.
 SQ SEQUENCE 403 AA; 45741 MW; 907305 CN;

Query Match 2.8%; Score 147; DB 7; Length 403;
 Best Local Similarity 28.7%; Pred. No. 6.42e-08;
 Matches 66; Conservative 49; Mismatches 84; Indels 31;
 Gaps 22;

```

D               b               6               1
nyittsikvedadkdtlvcgtnnngnpkcwk--i-dgs--ddpkhrgrgyapyqnskvtti 115
      |||  |:| | : | :|||  :| | :  | |  :  |||  |::
:
Q               y               1               0               9
NYIRVLAKIDD-DR-VLICGTNAYKPLCRHYALKDGDYVVEKEYEGRGLCPFDPDHNSTA 166

D               b               1               1               6
synecvlsdinisk-egikrwr rfdgpcgydlytadnvipkdglrga-fvdkdgydkvy 173
      |:| |  ::  |  : ||  |  ::  |: | :: ||:
| ::
Q               y               1               6               7
IYSEGQLYSATVADFS GTDPLI-YRGP----L-RTERS DLKQ-LNAPNFVNTMEYND FIF 219

D               b               1               7               4
ilftdt-igskrivkipy--iaqmclndeggpsslsshrwstflkvelecdidgr-s-yr 228
      ::| :| :  |  |  :|::| :| |||  :: ||::|||  |:| : |
: |
Q               y               2               2               0
FFFRETAVEYINCGKAIYSRVARVCKHDKGGPHQ-GGDRWTSFLKSRLNCSVPGDYPFYF 278

Db      229 qiihsrt--iktd-ndti--l-yvffdsp-y sk--salctysmntikqsf 269
      |:| : | :  :  | | | :| |  ||::||::||::| :||
Qy      279 NEIQSTSDIIEGNYGGQVEKLIYGVFTTPVNSIGGS AVCAFSMK SILESF 328
  
```

Search completed: Thu May 16 15:23:58 1996
 Job time : 49 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:29:05 1996; MasPar time 9.00
Seconds

598.003 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-60

Description: (1-650) from US08121713B.pep

Perfect Score: 4680

S e q u e n c e :

EDDCQNYIRIMVVPSPGRLF.....TTPSSRIVVATTSEHCVPTR 650

1

Scoring table: PAM 150
Gap 11

Searched: 70887 seqs, 8282111 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8 9:part9 10:part10 11:part11 12:part12
13:part13
14:part14

Statistics: Mean 36.494; Variance 155.724; scale 0.234

Pred. No. is the number of results predicted by chance to
have a

score greater than or equal to the score of the result being
printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description
Pred. No.						

1	4680	100.0	650	13	R71382	Drosophila semaphorin

0.00e+00	2	2060	44.0	730	13	R71379	Grasshopper semaphori
1.87e-189	3	1595	34.1	712	13	R71384	Tribolium semaphorin
5.76e-143	4	898	19.2	724	13	R71383	Drosophila semaphorin
6.63e-74	5	804	17.2	771	13	R71380	Human semaphorin III
1.07e-64	6	614	13.1	477	13	R74175	Human collapsin.
3.07e-46	7	125	2.7	441	13	R71381	Vaccinia virus semaph
1.36e-01	8	99	2.1	216	5	R27223	GTP binding protein.
1.13e+01	9	98	2.1	685	14	R77781	Polyphosphate kinase.
1.33e+01	10	98	2.1	820	11	R53228	Rice starch branching
1.33e+01	11	98	2.1	820	9	R47468	Branching enzyme of r
1.33e+01	12	95	2.0	122	13	R71385	Variola major virus s
2.15e+01	13	94	2.0	751	4	R23582	Branching enzyme.
2.52e+01	14	94	2.0	768	5	R27683	Rabbit beta-8 intergi
2.52e+01	15	92	2.0	1195	1	R05530	High density lipoprot
3.45e+01	16	92	2.0	1292	1	R05531	High density lipoprot
3.45e+01	17	95	2.0	3079	11	R59926	GAP protein Ira2.
2.15e+01	18	89	1.9	215	13	R77287	T-cell receptor alpha
5.51e+01	19	87	1.9	226	3	R15661	C-terminally truncate
7.50e+01	20	87	1.9	246	3	R15660	Osmotin-like antifung
7.50e+01	21	88	1.9	250	3	P60726	Triose phosphate isom
6.43e+01	22	89	1.9	268	3	P60237	Sequence of the alpha
5.51e+01	23	89	1.9	268	13	R77288	T-cell receptor alpha
5.51e+01	24	89	1.9	270	3	P50256	Sequence of T-cell an
5.51e+01	25	90	1.9	393	12	R60700	Human ERK-5.
4.72e+01	26	89	1.9	547	12	R67381	L. pneumophila HtpB g
5.51e+01	27	89	1.9	667	5	R26912	ALS C3 mutant of toba
5.51e+01	28	89	1.9	667	3	R11974	Tobacco SURA-C3 mutan

5.51e+01	29	89	1.9	667	1	P81150	Herbicide resistant f
5.51e+01	30	89	1.9	667	12	R63081	ALS SURA-C3 (P197E) m
5.51e+01	31	89	1.9	768	5	R27684	Human beta-8 intergin
5.51e+01	32	91	1.9	783	13	R70841	E. coli polymerase-II
4.04e+01	33	91	1.9	783	5	R24441	E. coli DNA polymeras
4.04e+01	34	89	1.9	950	2	R06376	Product of the sscl g
5.51e+01	35	88	1.9	1822	10	R55273	Beta subunit of integ
6.43e+01	36	84	1.8	310	2	R11119	Recombinant L-2-hydro
1.18e+02	37	84	1.8	334	3	R21413	NADH dehydrogenase 1.
1.18e+02	38	84	1.8	491	13	R72369	Human auxillary cytoc
1.18e+02	39	85	1.8	657	6	R28964	Notch hN5k full lengt
1.02e+02	40	85	1.8	740	5	R27530	Plasmodium falciparum
1.02e+02	41	85	1.8	740	13	R68838	Plasmodium falciparum
1.02e+02	42	84	1.8	825	12	R60811	Rice starch branching
1.18e+02	43	84	1.8	958	1	P81181	Sequence of glucoamyl
1.18e+02	44	84	1.8	958	2	R07575	Glucoamylase encoded
1.18e+02	45	85	1.8	2938	11	R59923	GAP protein Iral.
1.02e+02							

ALIGNMENTS

RESULT 1
 ID R71382 standard; Protein; 650 AA.
 AC R71382;
 DT 21-NOV-1995 (first entry)
 DE Drosophila semaphorin I protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
 Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding
 activity;
 KW modulation; nerve cell growth; immune response; viral
 pathogenesis;
 KW neurological disease; neuro-regeneration; oncological
 infection.
 OS Drosophila sp.
 PN WO9507706-A.

PD 23-MAR-1995.
 PF 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 DR WPI; 95-131177/17.
 DR N-PSDB; Q87444.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2; Page 74-77; 101pp; English.
 CC The sequence of the Drosophila semaphorin I protein. The gene
 was
 CC isolated by PCR using primers based on sequence homology
 between the
 CC complete grasshopper and partial Tribolium and Manduca (moth)
 semaphorin
 CC gene sequences. The products gave two different sequences,
 each of which
 CC was used to obtain its respective complete sequence:
 semaphorin I and II
 CC (Q87445). The proteins encoded by the grasshopper semaphorin
 I (Q87441),
 CC human semaphorin III (Q87442), vaccinia virus semaphorin IV,
 Drosophila
 CC semaphorin I and II, Tribolium semaphorin I (Q87446) or
 variola major
 CC (smallpox) virus semaphorin IV (Q87447) genes were used to
 generate a
 CC series of peptides (R70370-R70418), which retain semaphorin
 receptor
 CC binding activity. The semaphorin derived or semaphorin
 receptor derived
 CC peptides are potent modulators of nerve cell growth, immune
 CC responsiveness and viral pathogenesis. They can be used in
 diagnosis and
 CC treatment of neurological disease and neuro-regeneration,
 immune
 CC modulation and diagnosis and treatment of viral and
 oncological infection
 CC and diseases.
 SQ Sequence 650 AA;

Query Match 100.0%; Score 4680; DB 13; Length 650;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 650; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 Search completed: Thu May 16 15:29:43 1996
 Job time : 38 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:27:37 1996; MasPar time 19.47
Seconds

843.866 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-60

Description: (1-650) from US08121713B.pep

Perfect Score: 4680

S e q u e n c e :

EDDCQNYIRIMVVPSPGRLF.....TTPSSRIVVATTSEHCVPTR 650

1

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3
7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 49.809; Variance 113.161; scale 0.440

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
Pred. No.							
1	4600	98.3	656	10	B49423		semaphorin I - fruit
0.00e+00							

2	2077	44.4	730	10	JH0798	fasciclin IV precurs
0.00e+00						
3	1606	34.3	711	13	A49423	semaphorin I precurs
8.68e-263						
4	898	19.2	724	10	C49423	semaphorin II precu
1.02e-132						
5	815	17.4	772	10	A49069	collapsin - chicken
9.89e-118						
6	804	17.2	771	11	D49423	semaphorin III precu
9.44e-116						
7	139	3.0	329	8	H64115	phenylalanine--tRNA
1.98e-04						
8	125	2.7	295	6	JQ1775	SalL9R protein - vac
1.08e-02						
9	125	2.7	403	6	E42521	A39R protein - vacci
1.08e-02						
10	125	2.7	441	6	S29921	hypothetical protein
1.08e-02						
11	114	2.4	208	5	A38202	ypt family - maize
2.10e-01						
12	111	2.4	305	6	S08312	gamma-hordein 1 prec
4.59e-01						
13	114	2.4	331	1	SYECFA	phenylalanine--tRNA
2.10e-01						
14	107	2.3	201	5	S36187	GTP-binding protein
1.27e+00						
15	109	2.3	688	8	A47705	triacylglycerol lipa
7.65e-01						
16	107	2.3	1238	7	S17944	virulence protein bv
1.27e+00						
17	101	2.2	379	10	S31719	proline-rich protein
5.56e+00						
18	105	2.2	491	4	A31047	testosterone 16alpha
2.09e+00						
19	103	2.2	602	6	S38111	amino acid transport
3.42e+00						
20	103	2.2	757	6	A48592	transferrin receptor
3.42e+00						
21	104	2.2	792	7	S20554	pyruvate,water dikin
2.68e+00						
22	102	2.2	857	5	S05943	gelation factor - sl
4.36e+00						
23	102	2.2	971	3	JQ1634	outer capsid protein
4.36e+00						
24	103	2.2	1238	7	S17946	virulence protein bv
3.42e+00						
25	105	2.2	1518	8	A44811	glucosyltransferase
2.09e+00						
26	100	2.1	142	6	JQ1776	SalF1R protein - vac
7.06e+00						
27	96	2.1	203	13	S41789	hypothetical protein
1.81e+01						
28	96	2.1	216	5	JS0640	GTP-binding protein
1.81e+01						

29	96	2.1	216	4	S06620	DNA-directed RNA pol
1.81e+01						
30	96	2.1	216	5	S33900	GTP-binding protein
1.81e+01						
31	97	2.1	257	1	PWSPD	H ⁺ -transporting ATP
1.43e+01						
32	97	2.1	334	1	DEQYG	glyceraldehyde-3-pho
1.43e+01						
33	99	2.1	365	7	B30930	flagellar basal body
8.96e+00						
34	96	2.1	386	12	S47163	DNA-directed RNA pol
1.81e+01						
35	98	2.1	492	4	S27160	cytochrome P450 2B12
1.13e+01						
36	100	2.1	515	10	C40581	embryonic nuclear pr
7.06e+00						
37	100	2.1	537	10	B40581	embryonic nuclear pr
7.06e+00						
38	100	2.1	539	10	A40581	embryonic nuclear pr
7.06e+00						
39	96	2.1	605	12	S16561	glutamine--fructose-
1.81e+01						
40	100	2.1	622	11	S24241	protein kinase rck (
7.06e+00						
41	98	2.1	820	9	JX0243	1,4-alpha-glucan bra
1.13e+01						
42	98	2.1	820	9	S34037	1,4-alpha-glucan bra
1.13e+01						
43	100	2.1	1187	13	JC4155	PEZ protein - human
7.06e+00						
44	100	2.1	1189	11	JC2366	protein tyrosine pho
7.06e+00						
45	96	2.1	1238	7	A40185	virulence protein bv
1.81e+01						

ALIGNMENTS

```

RESULT      1
ENTRY       B49423      #type fragment
TITLE       semaphorin I - fruit fly (Drosophila melanogaster)
(fragment)
ORGANISM    #formal_name Drosophila melanogaster
DATE        06-Jan-1995 #sequence_revision 06-Jan-1995
#text_change
            27-Jan-1995
ACCESSIONS  B49423
REFERENCE   A49423
            #authors    Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
            #journal     Cell (1993) 75:1389-1399
            #title       The Semaphorin genes encode a family of
transmembrane and secreted growth cone guidance molecules.
            #accession   B49423

```

##status preliminary; not compared with conceptual
translation
##molecule_type mRNA
##residues 1-656 ##label KOL
##cross-references GB:L26082
##note nucleotide sequence not given

GENETICS

#gene semaI
SUMMARY #length 656 #checksum 3860

Query Match 98.3%; Score 4600; DB 10; Length 656;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 647; Conservative 2; Mismatches 1; Indels 6;
Gaps 1;
Search completed: Thu May 16 15:28:48 1996
Job time : 71 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:26:40 1996; MasPar time 11.43
Seconds

871.717 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-60

Description: (1-650) from US08121713B.pep

Perfect Score: 4680

S e q u e n c e :

EDDCQNYIRIMVVPSPGRLF.....TTPSSRIVVATTSEHCVPTR 650

1

Scoring table: PAM 150
Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot31
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8

Statistics: Mean 51.745; Variance 90.696; scale 0.571

Pred. No. is the number of results predicted by chance to
have a

score greater than or equal to the score of the result being
printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
Pred. No.							
1	125	2.7	403	7	VA39_VACCC	PROTEIN A39.	
3.11e-04							
2	125	2.7	441	7	VA39_VACCV	PROTEIN A39.	
3.11e-04							
3	114	2.4	208	8	YPT1_MAIZE	GTP-BINDING PROTEIN Y	

1.27e-02	4	111	2.4	305	4	HORG_HORVU	GAMMA-HORDEIN 1 PRECU
3.35e-02	5	114	2.4	327	7	SYFA_ECOLI	PHENYLALANYL-TRNA SYN
1.27e-02	6	107	2.3	201	6	RAB9_CANFA	RAS-RELATED PROTEIN R
1.19e-01	7	109	2.3	688	4	LIP_STAEP	LIPASE PRECURSOR (EC
6.33e-02	8	106	2.3	793	6	PPSA_ECOLI	PHOSPHOENOLPYRUVATE S
1.62e-01	9	107	2.3	1238	1	BVGS_BORBR	VIRULENCE SENSOR PROT
1.19e-01	10	105	2.2	491	2	CPB9_MOUSE	CYTOCHROME P450 IIB9
2.20e-01	11	103	2.2	602	3	GAP1_YEAST	GENERAL AMINO ACID PE
4.04e-01	12	102	2.2	857	3	GELA_DICDI	GELATION FACTOR (ACTI
5.46e-01	13	102	2.2	971	8	VP2_EHDV1	OUTER CAPSID PROTEIN
5.46e-01	14	103	2.2	1238	1	BVGS_BORPA	VIRULENCE SENSOR PROT
4.04e-01	15	96	2.1	203	8	YOR5_THER8	HYPOTHETICAL 23.9 KD
3.14e+00	16	96	2.1	216	6	RPOU_METVA	DNA-DIRECTED RNA POLY
3.14e+00	17	96	2.1	216	1	ARA3_ARATH	RAS-RELATED PROTEIN A
3.14e+00	18	97	2.1	257	1	ATPD_SPIOL	ATP SYNTHASE DELTA CH
2.36e+00	19	97	2.1	334	3	G3P_PYRWO	GLYCERALDEHYDE 3-PHOS
2.36e+00	20	99	2.1	365	3	FLGI_SALTY	FLAGELLAR P-RING PROT
1.32e+00	21	98	2.1	492	2	CPBC_RAT	CYTOCHROME P450 IIB12
1.77e+00	22	96	2.1	604	5	NODM_RHIME	GLUCOSAMINE--FRUCTOSE
3.14e+00	23	100	2.1	622	4	KRCK_MOUSE	PROTEIN KINASE RCK (E
9.88e-01	24	98	2.1	684	6	PPK_KLEAE	POLYPHOSPHATE KINASE
1.77e+00	25	98	2.1	820	3	GLGB_ORYSA	1,4-ALPHA-GLUCAN BRAN
1.77e+00	26	96	2.1	1238	1	BVGS_BORPE	VIRULENCE SENSOR PROT
3.14e+00	27	93	2.0	195	1	ARA5_ARATH	RAS-RELATED PROTEIN A
7.25e+00	28	94	2.0	217	6	RIC2_ORYSA	RAS-RELATED PROTEIN R
5.50e+00	29	92	2.0	218	8	YPT3_NICPL	GTP-BINDING PROTEIN Y
9.53e+00	30	92	2.0	246	6	PYG2_ANASP	PHYCOBILISOME ROD-COR

9.53e+00	31	94	2.0	287	7	UL24_ILTVT	PROTEIN UL24 HOMOLOG
5.50e+00	32	92	2.0	330	5	MTHT_METTF	MODIFICATION METHYLAS
9.53e+00	33	93	2.0	357	5	MTLD_ENTFA	MANNITOL-1-PHOSPHATE
7.25e+00	34	92	2.0	380	3	GAL1_SALTY	GALACTOKINASE (EC 2.7
9.53e+00	35	94	2.0	649	7	SYM_BACST	METHIONYL-TRNA SYNTH
5.50e+00	36	95	2.0	687	6	PPK_ECOLI	POLYPHOSPHATE KINASE
4.16e+00	37	94	2.0	768	4	ITB8_RABIT	INTEGRIN BETA-8 SUBUN
5.50e+00	38	95	2.0	783	8	YAF3_YEAST	HYPOTHETICAL 87.5 KD
4.16e+00	39	94	2.0	825	7	SWI3_YEAST	TRANSCRIPTION REGULAT
5.50e+00	40	92	2.0	1004	5	MV10_MOUSE	PUTATIVE GTP-BINDING
9.53e+00	41	93	2.0	1041	6	RAG1_CHICK	V(D)J RECOMBINATION A
7.25e+00	42	95	2.0	1379	5	MET_MOUSE	HEPATOCYTE GROWTH FAC
4.16e+00	43	94	2.0	1390	5	MET_HUMAN	HEPATOCYTE GROWTH FAC
5.50e+00	44	95	2.0	2366	7	TOXB_CLODI	TOXIN B.
4.16e+00	45	95	2.0	3079	4	IRA2_YEAST	INHIBITORY REGULATOR
4.16e+00							

ALIGNMENTS

RESULT 1
 ID VA39 VACCC STANDARD; PRT; 403 AA.
 AC P21062;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE PROTEIN A39.
 GN A39R.
 OS VACCINIA VIRUS (STRAIN COPENHAGEN).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE;
 CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91021027
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW
 J.P.,
 RA PAOLETTI E.;
 RL VIROLOGY 179:247-266(1990).

RN [2]
 RP COMPLETE GENOME.
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
 RA PAOLETTI E.;
 RL VIROLOGY 179:517-563(1990).
 DR EMBL; M35027; PXVACCG.
 DR PIR; E42521; E42521.
 SQ SEQUENCE 403 AA; 45741 MW; 907305 CN;

Query Match 2.7%; Score 125; DB 7; Length 403;
 Best Local Similarity 37.3%; Pred. No. 3.11e-04;
 Matches 22; Conservative 12; Mismatches 20; Indels 5;
 Gaps 4;

D	b	1	7	2
vyilftdt-igskrivkipy--iaqmcIndeggpsslsshrwtflkvelecdidgrsy				227
	:: : :	: :: :	: : :: :	
Q	y	1	1	5
VYFFFRRETAVEFINCGKAIYSRVARVCKWDKGGPHRFR-NRWTSFLKSRLNCSIPG-DY				171

Search completed: Thu May 16 15:27:20 1996
 Job time : 40 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:32:28 1996; MasPar time 9.80
Seconds

611.795 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-62

Description: (1-724) from US08121713B.pep

Perfect Score: 5375

S e q u e n c e :

MSLLQLSPLLALLLLLCSSV.....TRQNFSCNQHPNEIFRKPNV 724

1

Scoring table: PAM 150
Gap 11

Searched: 70887 seqs, 8282111 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8 9:part9 10:part10 11:part11 12:part12
13:part13
14:part14

Statistics: Mean 36.943; Variance 161.422; scale 0.229

Pred. No. is the number of results predicted by chance to
have a

score greater than or equal to the score of the result being
printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description
Pred. No.						

1	5375	100.0	724	13	R71383	Drosophila semaphorin

0.00e+00	2	963	17.9	730	13	R71379	Grasshopper semaphori
3.25e-78	3	898	16.7	650	13	R71382	Drosophila semaphorin
5.59e-72	4	886	16.5	712	13	R71384	Tribolium semaphorin
7.89e-71	5	799	14.9	771	13	R71380	Human semaphorin III
1.63e-62	6	646	12.0	477	13	R74175	Human collapsin.
5.43e-48	7	150	2.8	441	13	R71381	Vaccinia virus semaph
2.43e-03	8	109	2.0	122	13	R71385	Variola major virus s
2.95e+00	9	105	2.0	1257	9	R46627	Neurocan core protein
5.65e+00	10	95	1.8	124	11	R59943	Anti-VLA4 rAb humanis
2.76e+01	11	95	1.8	235	14	R70101	LIYV RNA 2 ORF 6 pept
2.76e+01	12	99	1.8	236	1	P93910	Y22 light chain of an
1.47e+01	13	99	1.8	371	1	P93558	Linkered human IgG1 (
1.47e+01	14	95	1.8	371	1	P91918	Sequence of the linke
2.76e+01	15	95	1.8	452	14	R70420	Lettuce infectious ye
2.76e+01	16	96	1.8	554	13	R70827	Anti-cataract immunot
2.36e+01	17	96	1.8	1451	5	R27819	CCVInsavc spike prote
2.36e+01	18	97	1.8	1822	10	R55273	Beta subunit of integ
2.02e+01	19	90	1.7	94	4	R22580	Light chain VL16.1 fr
5.93e+01	20	91	1.7	106	8	R39818	HP1/2 Vk.
5.09e+01	21	91	1.7	106	11	R59937	HP1/2 Vk.
5.09e+01	22	91	1.7	106	8	R39567	Sequence of the varia
5.09e+01	23	91	1.7	106	11	R58750	Anti-VLA4 Ab HP1/2 V-
5.09e+01	24	89	1.7	107	5	R25413	Light chain variable
6.89e+01	25	91	1.7	128	11	R59935	VK3 (DQMDY) VL.
5.09e+01	26	91	1.7	128	11	R58752	Anti-VLA4 SVM DY V-kap
5.09e+01	27	91	1.7	128	11	R59934	VK2 (SVM DY) VL.
5.09e+01	28	91	1.7	129	13	R70256	Anti-VLA-4 humanized

5.09e+01	29	92	1.7	154	13	R70828	MAb 4197X light chain
4.37e+01	30	93	1.7	278	3	P60007	Sequence of oncogene
3.75e+01	31	90	1.7	304	1	P91816	Mammalian T lymphocyt
5.93e+01	32	89	1.7	351	8	R41542	B15R product.
6.89e+01	33	89	1.7	351	2	R24251	Vaccinia virus B18R p
6.89e+01	34	89	1.7	354	2	R06429	SP1-like protein enco
6.89e+01	35	89	1.7	375	14	R79120	Neuropeptide Y/peptid
6.89e+01	36	90	1.7	407	13	R65494	Marek's disease virus
5.93e+01	37	91	1.7	419	2	R06428	SP1-like protein enco
5.09e+01	38	94	1.7	424	2	R06430	SP1-like protein enco
3.22e+01	39	94	1.7	424	2	P93998	Amino acid sequence o
3.22e+01	40	91	1.7	426	3	R12458	Pregnancy-specific be
5.09e+01	41	93	1.7	699	3	R14202	Gamma-cyclodextrin gl
3.75e+01	42	91	1.7	830	13	R65216	P-selectin.
5.09e+01	43	88	1.6	489	8	R42176	Murine MDM2.
8.01e+01	44	88	1.6	750	1	P82867	Enkephalinase (rat).
8.01e+01	45	88	1.6	750	1	P90393	Human common acute ly
8.01e+01							

ALIGNMENTS

RESULT 1
 ID R71383 standard; Protein; 724 AA.
 AC R71383;
 DT 21-NOV-1995 (first entry)
 DE Drosophila semaphorin II protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
 Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding
 activity;
 KW modulation; nerve cell growth; immune response; viral
 pathogenesis;
 KW neurological disease; neuro-regeneration; oncological
 infection.
 OS Drosophila sp.
 PN WO9507706-A.

PD 23-MAR-1995.
 PF 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 DR WPI; 95-131177/17.
 DR N-PSDB; Q87445.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2; Page 79-82; 101pp; English.
 CC The sequence of the Drosophila semaphorin II protein. The
 gene was
 CC isolated by PCR using primers based on sequence homology
 between the
 CC complete grasshopper and partial Tribolium and Manduca (moth)
 semaphorin
 CC gene sequences. The products gave two different sequences,
 each of which
 CC was used to obtain its respective complete sequence:
 semaphorin I
 CC (Q87444) and II. The proteins encoded by the grasshopper
 semaphorin I
 CC (Q87441), human semaphorin III (Q87442), vaccinia virus
 semaphorin IV,
 CC Drosophila semaphorin I and II, Tribolium semaphorin I
 (Q87446) or
 CC variola major (smallpox) virus semaphorin IV (Q87447) genes
 were used to
 CC generate a series of peptides (R70370-R70418), which retain
 semaphorin
 CC receptor binding activity. The semaphorin derived or
 semaphorin receptor
 CC derived peptides are potent modulators of nerve cell growth,
 immune
 CC responsiveness and viral pathogenesis. They can be used in
 diagnosis and
 CC treatment of neurological disease and neuro-regeneration,
 immune
 CC modulation and diagnosis and treatment of viral and
 oncological infection
 CC and diseases.
 SQ Sequence 724 AA;

Query Match 100.0%; Score 5375; DB 13; Length 724;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 724; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 Search completed: Thu May 16 15:33:08 1996
 Job time : 40 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:31:05 1996; MasPar time 21.38
Seconds

855.707 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-62

Description: (1-724) from US08121713B.pep

Perfect Score: 5375

S e q u e n c e :

MSLLQLSPLLALLLLLCSSV.....TRQNFSCNQHPNEIFRKPNV 724

1

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3
7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 49.968; Variance 111.741; scale 0.447

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
No.						
Pred. No.						

1	5375	100.0	724	10	C49423	semaphorin II precur
0.00e+00						

2	970	18.0	730	10	JH0798	fasciclin IV precurs	
4.78e-148	3	898	16.7	656	10	B49423	semaphorin I - fruit
8.89e-135	4	897	16.7	711	13	A49423	semaphorin I precurs
1.36e-134	5	806	15.0	772	10	A49069	collapsin - chicken
6.56e-118	6	799	14.9	771	11	D49423	semaphorin III precu
1.25e-116	7	150	2.8	403	6	E42521	A39R protein - vacci
5.83e-06	8	150	2.8	441	6	S29921	hypothetical protein
5.83e-06	9	133	2.5	295	6	JQ1775	SalL9R protein - vac
9.48e-04	10	115	2.1	466	1	DCBYO	ornithine decarboxyl
1.45e-01	11	112	2.1	637	1	HJECDR	helicase (EC 3.6.1.-
3.21e-01	12	114	2.1	1149	4	S14169	DNA-directed RNA pol
1.89e-01	13	114	2.1	1287	4	A43488	genome polyprotein -
1.89e-01	14	115	2.1	1535	13	S46224	peroxidasin - fruit
1.45e-01	15	112	2.1	1763	1	RRWWF9	RNA-directed RNA pol
3.21e-01	16	109	2.0	122	6	H36852	A43R protein - vario
7.00e-01	17	109	2.0	122	6	JQ1845	14R protein - variol
7.00e-01	18	107	2.0	240	13	JC4121	pregnancy-specific g
1.17e+00	19	107	2.0	1087	12	S41797	xylanase (EC 3.2.1.9
1.17e+00	20	105	2.0	1257	11	S28764	neurocan - rat
1.94e+00	21	103	1.9	253	8	JQ0091	hypothetical 29K pro
3.19e+00	22	104	1.9	319	7	S49263	shiga-like toxin typ
2.49e+00	23	103	1.9	339	9	S45471	hypothetical protein
3.19e+00	24	103	1.9	379	12	S26851	site-specific DNA-me
3.19e+00	25	100	1.9	664	10	S44756	C14B9.2 protein - Ca
6.64e+00	26	100	1.9	1147	11	A41674	myosin-light-chain k
6.64e+00	27	100	1.9	1176	11	JN0583	myosin-light-chain k
6.64e+00	28	104	1.9	1268	11	S52781	neurocan - mouse
2.49e+00							

29	103	1.9	1314	9	S19488	probable membrane pr
3.19e+00						
30	103	1.9	1390	1	TVHUME	hepatocyte growth fa
3.19e+00						
31	101	1.9	1404	10	A48196	transforming protein
5.21e+00						
32	99	1.8	111	5	PH1028	Ig heavy chain V reg
8.45e+00						
33	98	1.8	206	6	S31681	hypothetical protein
1.07e+01						
34	97	1.8	452	6	A36429	integrin beta-4 chai
1.36e+01						
35	98	1.8	524	10	S35341	kettin - fruit fly (
1.07e+01						
36	98	1.8	769	9	S50966	hypothetical protein
1.07e+01						
37	98	1.8	890	3	JQ1947	genome polyprotein 2
1.07e+01						
38	99	1.8	1323	10	PN0568	connectin 3B - chick
8.45e+00						
39	97	1.8	1805	13	S12380	integrin beta-4 chai
1.36e+01						
40	99	1.8	1951	10	S27356	aggrecan - chicken
8.45e+00						
41	99	1.8	2071	13	S39796	aggrecan - chicken
8.45e+00						
42	98	1.8	2292	3	GNNYED	genome polyprotein -
1.07e+01						
43	98	1.8	2292	3	GNNYEB	genome polyprotein -
1.07e+01						
44	98	1.8	2292	12	S55401	capsid polyprotein p
1.07e+01						
45	97	1.8	2292	6	S35961	capsid polyprotein p
1.36e+01						

ALIGNMENTS

```

RESULT      1
ENTRY       C49423      #type complete
TITLE       semaphorin II precursor - fruit fly (Drosophila
melanogaster)
ORGANISM    #formal_name Drosophila melanogaster
DATE        06-Jan-1995 #sequence_revision 06-Jan-1995
#text_change
            27-Jan-1995
ACCESSIONS  C49423
REFERENCE   A49423
            #authors   Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
            #journal    Cell (1993) 75:1389-1399
            #title      The Semaphorin genes encode a family of
transmembrane and      secreted growth cone guidance molecules.
            #accession  C49423

```

##status preliminary; not compared with conceptual
translation
##molecule_type mRNA
##residues 1-724 ##label KOL
##cross-references GB:L26083
##note nucleotide sequence not given

GENETICS

#gene sema II
SUMMARY #length 724 #molecular-weight 82971 #checksum
6082

Query Match 100.0%; Score 5375; DB 10; Length 724;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 724; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Search completed: Thu May 16 15:32:11 1996
Job time : 66 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:30:00 1996; MasPar time 12.58
Seconds

882.268 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-62

Description: (1-724) from US08121713B.pep

Perfect Score: 5375

S e q u e n c e :

MSLLQLSPLLALLLLLCSSV.....TRQNFSCNQHPNEIFRKPNV 724

1

Scoring table: PAM 150
Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot31
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8

Statistics: Mean 51.877; Variance 88.790; scale 0.584

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
Pred. No.							
1	150	2.8	403	7	VA39_VACCC	PROTEIN A39.	
1.94e-08							
2	150	2.8	441	7	VA39_VACCV	PROTEIN A39.	

1.94e-08	3	115	2.1	466	2	DCOR_YEAST	ORNITHINE DECARBOXYLA
7.14e-03	4	114	2.1	1149	6	RPC2_YEAST	DNA-DIRECTED RNA POLY
9.98e-03	5	114	2.1	1287	6	POLN_FCVC6	NON-STRUCTURAL POLYPR
9.98e-03	6	112	2.1	1763	6	POLN_FCVF9	NON-STRUCTURAL POLYPR
1.94e-02	7	107	2.0	1087	8	XYNX_CLOTM	EXOGLUCANASE XYNX PRE
9.77e-02	8	103	1.9	253	8	YMG1_MYCGE	HYPOTHETICAL 29.1 KD
3.42e-01	9	103	1.9	339	8	YBS6_YEAST	HYPOTHETICAL 40.2 KD
3.42e-01	10	100	1.9	664	3	ER72_CAEEL	PROBABLE ERP72 PROTEI
8.50e-01	11	103	1.9	1314	4	KCW3_YEAST	PUTATIVE SERINE/THREO
3.42e-01	12	103	1.9	1390	5	MET_HUMAN	HEPATOCYTE GROWTH FAC
3.42e-01	13	95	1.8	115	7	TVC2_MOUSE	T-CELL RECEPTOR GAMMA
3.68e+00	14	95	1.8	135	7	TVC4_MOUSE	T-CELL RECEPTOR GAMMA
3.68e+00	15	96	1.8	492	8	VSM1_TRYBB	VARIANT SURFACE GLYCO
2.76e+00	16	95	1.8	575	6	PUT2_YEAST	DELTA-1-PYRROLINE-5-C
3.68e+00	17	96	1.8	647	7	VE1_HPV34	E1 PROTEIN.
2.76e+00	18	95	1.8	659	1	BYR2_SCHPO	PROTEIN KINASE BYR2 (
3.68e+00	19	96	1.8	737	7	VE02_VARV	PROTEIN E2.
2.76e+00	20	98	1.8	890	6	POL2_BAYMG	GENOME POLYPROTEIN 2
1.54e+00	21	97	1.8	1875	4	ITB4_HUMAN	INTEGRIN BETA-4 SUBUN
2.07e+00	22	97	1.8	2290	6	POLG_EMCV	GENOME POLYPROTEIN (C
2.07e+00	23	98	1.8	2292	6	POLG_EMCVD	GENOME POLYPROTEIN (C
1.54e+00	24	98	1.8	2292	6	POLG_EMCVB	GENOME POLYPROTEIN (C
1.54e+00	25	96	1.8	2303	6	POLG_TMEVG	GENOME POLYPROTEIN (C
2.76e+00	26	95	1.8	2303	6	POLG_TMEVB	GENOME POLYPROTEIN (C
3.68e+00	27	94	1.7	282	7	STR_STAAU	STREPTOMYCIN RESISTAN
4.90e+00	28	91	1.7	308	1	ABCA_AERSA	ABCA PROTEIN.
1.13e+01	29	92	1.7	318	7	SLTA_BP933	SHIGA-LIKE TOXIN II S

8.57e+00	30	94	1.7	326	8	VS09_ROTHV	GLYCOPROTEIN VP7 (SER
4.90e+00	31	93	1.7	368	8	YCZ2_YEAST	HYPOTHETICAL 40.1 KD
6.49e+00	32	91	1.7	393	3	FDH_PSESR	FORMATE DEHYDROGENASE
1.13e+01	33	91	1.7	417	5	PBG1_HUMAN	PREGNANCY-SPECIFIC BE
1.13e+01	34	91	1.7	419	5	PBGD_HUMAN	PREGNANCY-SPECIFIC BE
1.13e+01	35	91	1.7	428	5	PBGC_HUMAN	PREGNANCY-SPECIFIC BE
1.13e+01	36	94	1.7	556	2	CD19_HUMAN	B-LYMPHOCYTE ANTIGEN
4.90e+00	37	91	1.7	648	4	KMLC_CHICK	MYOSIN LIGHT CHAIN KI
1.13e+01	38	92	1.7	657	3	HCYB_PANIN	HEMOCYANIN B CHAIN.
8.57e+00	39	93	1.7	811	3	FS22_DROME	FASCICLIN II, PHOSPHA
6.49e+00	40	91	1.7	830	4	LEM3_HUMAN	P-SELECTIN PRECURSOR
1.13e+01	41	93	1.7	873	3	FS21_DROME	FASCICLIN II, MEMBRAN
6.49e+00	42	92	1.7	896	6	RPOP_NEUCR	PROBABLE DNA-DIRECTED
8.57e+00	43	91	1.7	1258	4	KML2_CHICK	MYOSIN LIGHT CHAIN KI
1.13e+01	44	94	1.7	1322	7	SUS_DROME	SUPPRESSOR OF SABLE P
4.90e+00	45	94	1.7	1451	8	VGL2_CVCAI	E2 GLYCOPROTEIN PRECU
4.90e+00							

ALIGNMENTS

RESULT 1
 ID VA39 VACCC STANDARD; PRT; 403 AA.
 AC P21062;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE PROTEIN A39.
 GN A39R.
 OS VACCINIA VIRUS (STRAIN COPENHAGEN).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE;
 CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91021027
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW
 J.P.,

RA PAOLETTI E.;
 RL VIROLOGY 179:247-266(1990).
 RN [2]
 RP COMPLETE GENOME.
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW
 J.P.,
 RA PAOLETTI E.;
 RL VIROLOGY 179:517-563(1990).
 DR EMBL; M35027; PXVACCG.
 DR PIR; E42521; E42521.
 SQ SEQUENCE 403 AA; 45741 MW; 907305 CN;

Query Match 2.8%; Score 150; DB 7; Length 403;
 Best Local Similarity 31.8%; Pred. No. 1.94e-08;
 Matches 34; Conservative 26; Mismatches 38; Indels 9;
 Gaps 8;

D	b	1	7	2
vyilftdt-igskrivkipy--iaqmclndeggpsslsshrwtflkvelecdidgr-s-	226			
: : : : : : : : : :				
:				
Q	y	2	6	5
VYFFFRETAVEYINCGKAVYSRIARVCKKDVGG-KNLLAHNWATYLNKARLNCSISGEFPF	323			
Db	227	yrqiihsr-tiktdndtilyvffdspys-k-salctysmntikqsfs	270	
Qy	324	YFNEIQSVYQLPSDKSRFFATFTTSTNGLIGSAVCSFHINEIQAAFN	370	

Search completed: Thu May 16 15:30:48 1996
 Job time : 48 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:36:00 1996; MasPar time 9.64
Seconds

611.897 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-64

Description: (1-712) from US08121713B.pep

Perfect Score: 5145

S e q u e n c e :

MVVKILVWSICLIALCHAWM.....SKDLNIASDGTLQIKKTYI 712

1

Scoring table: PAM 150
Gap 11

Searched: 70887 seqs, 8282111 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8 9:part9 10:part10 11:part11 12:part12
13:part13
14:part14

Statistics: Mean 36.742; Variance 151.200; scale 0.243

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	DB	ID	Description
Pred. No.						
1	5145	100.0	712	13	R71384	Tribolium semaphorin

0.00e+00	2	2394	46.5	730	13	R71379	Grasshopper semaphori
3.86e-231	3	1595	31.0	650	13	R71382	Drosophila semaphorin
3.26e-148	4	886	17.2	724	13	R71383	Drosophila semaphorin
2.09e-75	5	771	15.0	771	13	R71380	Human semaphorin III
9.58e-64	6	613	11.9	477	13	R74175	Human collapsin.
7.43e-48	7	144	2.8	441	13	R71381	Vaccinia virus semaph
2.90e-03	8	116	2.3	122	13	R71385	Variola major virus s
5.04e-01	9	97	1.9	700	3	R13227	Novel endoglucanase.
1.32e+01	10	96	1.9	763	10	R53778	Sequence of human amy
1.56e+01	11	99	1.9	1651	12	R63222	Pre-pro-cobra C3 prot
9.49e+00	12	93	1.8	261	14	R74184	Type I RIP pokeweed a
2.55e+01	13	93	1.8	261	3	R13112	Anti-viral protein.
2.55e+01	14	93	1.8	261	7	R37298	Plant type I RIP Poke
2.55e+01	15	92	1.8	1451	5	R27819	CCVInsavc spike prote
3.00e+01	16	86	1.7	158	14	R76778	Murine T-LIF.
7.81e+01	17	90	1.7	166	6	R32907	Vdelta3 (WM14).
4.14e+01	18	86	1.7	178	1	P82943	N-terminal of leukaem
7.81e+01	19	86	1.7	181	11	R60130	Mouse LIF.
7.81e+01	20	86	1.7	203	7	R33379	Cytokine mLIF.
7.81e+01	21	86	1.7	203	1	P82945	Leukaemia inhibitory
7.81e+01	22	85	1.7	216	4	P30205	Sequence encoded by d
9.13e+01	23	88	1.7	235	14	R70101	LIYV RNA 2 ORF 6 pept
5.70e+01	24	89	1.7	270	3	P60130	Murine IL-1 precursor
4.86e+01	25	88	1.7	294	7	R37345	Pokeweed antiviral pr
5.70e+01	26	89	1.7	331	8	R42260	Mature decorin PT-65.
4.86e+01	27	89	1.7	353	1	R05160	Sequence of human bon
4.86e+01	28	85	1.7	371	3	P60055	Partial Factor VII pe

9.13e+01	29	85	1.7	406	7	R35764	Factor VII (VII).
9.13e+01	30	85	1.7	406	4	P30203	Sequence encoded by p
9.13e+01	31	85	1.7	444	12	R64205	Factor VII - modified
9.13e+01	32	88	1.7	452	14	R70420	Lettuce infectious ye
5.70e+01	33	85	1.7	453	3	P60057	Factor IX/Factor VII
9.13e+01	34	85	1.7	466	3	P60056	Factor VII peptide en
9.13e+01	35	85	1.7	466	8	R52562	Factor VIII.
9.13e+01	36	85	1.7	1071	12	R60796	Rice sucrose phosphat
9.13e+01	37	85	1.7	1536	12	R63505	Haemophilus high mole
9.13e+01	38	85	1.7	1536	8	R41723	High molecular weight
9.13e+01	39	85	1.7	1536	8	R41725	High molecular weight
9.13e+01	40	89	1.7	1642	12	R63223	Cobra CVF1 .
4.86e+01	41	88	1.7	3144	11	R58777	Protein encoded by Hu
5.70e+01	42	87	1.7	3722	2	R10145	Cephalosporin antibio
6.67e+01	43	83	1.6	106	7	R35638	Tryptophan aporepress
1.24e+02	44	83	1.6	220	4	P30195	Sequence encoded by V
1.24e+02	45	84	1.6	1399	8	R38698	S-PRV-055 TGE virus g
1.07e+02							

ALIGNMENTS

RESULT 1
 ID R71384 standard; Protein; 712 AA.
 AC R71384;
 DT 21-NOV-1995 (first entry)
 DE Tribolium semaphorin I protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
 Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding
 activity;
 KW modulation; nerve cell growth; immune response; viral
 pathogenesis;
 KW neurological disease; neuro-regeneration; oncological
 infection.
 OS Tribolium sp.
 PN WO9507706-A.

PD 23-MAR-1995.
 PF 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 DR WPI; 95-131177/17.
 DR N-PSDB; Q87446.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2; Page 85-88; 101pp; English.
 CC The sequence of the beetle Tribolium semaphorin I protein.
 The gene was
 CC isolated by PCR using Tribolium genomic DNA. The proteins
 encoded by the
 CC grasshopper semaphorin I (Q87441), human semaphorin III
 (Q87442),
 CC vaccinia virus semaphorin IV (Q87443, Drosophila semaphorin I
 and II
 CC (Q87444-5), Tribolium semaphorin I or variola major (smallpox)
 virus
 CC semaphorin IV (Q87447) genes were used to generate a series of
 peptides
 CC (R70370-R70418), which retain semaphorin receptor binding
 activity. The
 CC semaphorin derived or semaphorin receptor derived peptides are
 potent
 CC modulators of nerve cell growth, immune responsiveness and
 viral
 CC pathogenesis. They can be used in diagnosis and treatment of
 neurological
 CC disease and neuro-regeneration, immune modulation and
 diagnosis and
 CC treatment of viral and oncological infection and diseases.
 SQ Sequence 712 AA;

Query Match 100.0%; Score 5145; DB 13; Length 712;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 712; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

Search completed: Thu May 16 15:36:34 1996
 Job time : 34 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:34:41 1996; MasPar time 21.12
Seconds

851.914 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-64

Description: (1-712) from US08121713B.pep

Perfect Score: 5145

S e q u e n c e :

MVVKILVWSICLIALCHAWM.....SKDLNIASDGTLQKIKKTYI 712

1

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3
7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 49.858; Variance 106.968; scale 0.466

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
Pred. No.							
1	5127	99.7	711	13	A49423		semaphorin I precurs
0.00e+00							

2	2389	46.4	730	10	JH0798	fasciclin IV precurs	
0.00e+00	3	1604	31.2	656	10	B49423	semaphorin I - fruit
5.91e-278	4	886	17.2	724	10	C49423	semaphorin II precur
2.46e-138	5	789	15.3	772	10	A49069	collapsin - chicken
8.12e-120	6	771	15.0	771	11	D49423	semaphorin III precu
2.14e-116	7	144	2.8	295	6	JQ1775	SalL9R protein - vac
1.55e-05	8	144	2.8	403	6	E42521	A39R protein - vacci
1.55e-05	9	144	2.8	441	6	S29921	hypothetical protein
1.55e-05	10	125	2.4	329	8	H64115	phenylalanine--tRNA
5.01e-03	11	116	2.3	122	6	JQ1845	14R protein - variol
6.61e-02	12	116	2.3	122	6	H36852	A43R protein - vario
6.61e-02	13	113	2.2	103	12	S45435	hypothetical protein
1.52e-01	14	109	2.1	331	1	SYECFA	phenylalanine--tRNA
4.51e-01	15	107	2.1	2769	2	UIBO	thyroglobulin precur
7.68e-01	16	105	2.0	195	6	B36838	D11L protein - vario
1.30e+00	17	104	2.0	642	1	SYRTAL	5-aminolevulinate sy
1.69e+00	18	102	2.0	663	9	S55164	hypothetical protein
2.82e+00	19	103	2.0	796	6	A35775	integrin beta-5 chai
2.18e+00	20	103	2.0	799	6	A38308	integrin beta-5 chai
2.18e+00	21	103	2.0	799	6	S12534	integrin beta-5 chai
2.18e+00	22	96	1.9	166	11	S04934	T-cell receptor delt
1.26e+01	23	98	1.9	191	5	A35981	sperm membrane prote
7.73e+00	24	97	1.9	197	7	S01240	heat shock protein B
9.89e+00	25	96	1.9	213	8	PQ0489	cyclin 1 - alfalfa (
1.26e+01	26	96	1.9	213	8	S29924	cyclin - alfalfa (fr
1.26e+01	27	98	1.9	315	3	WZVZB4	37K HindIII-C protei
7.73e+00	28	99	1.9	339	7	A42259	endo-beta-N-acetylgl
6.03e+00							

29	98	1.9	355	10	JH0446	75K autoantigen - hu
7.73e+00						
30	98	1.9	402	12	S55465	chalcone synthase 2
7.73e+00						
31	97	1.9	585	13	S55466	calicin - human (fra
9.89e+00						
32	99	1.9	689	12	S47780	glycine-tRNA ligase
6.03e+00						
33	99	1.9	689	1	SYECGB	glycine--tRNA ligase
6.03e+00						
34	97	1.9	700	8	B41897	cellulase (EC 3.2.1.
9.89e+00						
35	96	1.9	763	5	A49321	amyloid precursor pr
1.26e+01						
36	98	1.9	765	5	S42880	amyloid precursor-li
7.73e+00						
37	98	1.9	765	11	S47528	amyloid precursor-li
7.73e+00						
38	97	1.9	875	1	ITECAP	DNA topoisomerase (A
9.89e+00						
39	96	1.9	1133	6	S12597	M polyprotein precur
1.26e+01						
40	96	1.9	1133	3	GNVUSR	M polyprotein precur
1.26e+01						
41	96	1.9	1133	3	A43964	M polyprotein precur
1.26e+01						
42	96	1.9	1276	12	S11455	botulinum neurotoxin
1.26e+01						
43	96	1.9	1390	1	TVHUME	hepatocyte growth fa
1.26e+01						
44	96	1.9	1451	3	JQ1719	E2 glycoprotein prec
1.26e+01						
45	99	1.9	1651	2	C3NJ	complement C3 precur
6.03e+00						

ALIGNMENTS

```

RESULT      1
ENTRY       A49423      #type complete
TITLE       semaphorin I precursor - beetle (Tribolium
            confusum)
ORGANISM    #formal_name Tribolium confusum
DATE        12-Dec-1994 #sequence_revision 12-Dec-1994
#text_change
            12-Dec-1994
ACCESSIONS  A49423
REFERENCE   A49423
            #authors    Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
            #journal     Cell (1993) 75:1389-1399
            #title       The Semaphorin genes encode a family of
transmembrane and secreted growth cone guidance molecules.
            #accession   A49423

```

##status preliminary
##residues 1-711 ##label KOL
##cross-references GB:L26080
SUMMARY #length 711 #molecular-weight 79623 #checksum
4552

Query Match 99.7%; Score 5127; DB 13; Length 711;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 711; Conservative 0; Mismatches 0; Indels 1;
Gaps 1;
Search completed: Thu May 16 15:35:42 1996
Job time : 61 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:33:25 1996; MasPar time 12.38
Seconds

881.667 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-64

Description: (1-712) from US08121713B.pep

Perfect Score: 5145

S e q u e n c e :

MVVKILVWSICLIALCHAWM.....SKDLNIASDGT LQKIKKTYI 712

1

Scoring table: PAM 150

Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot31

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6

7:part7

8:part8

Statistics: Mean 51.891; Variance 83.088; scale 0.625

Pred. No. is the number of results predicted by chance to
have a

score greater than or equal to the score of the result being
printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
--------	-------	---------------------	--------	----	----	-------------

1	144	2.8	403	7	VA39_VACCC	PROTEIN A39.
2	144	2.8	441	7	VA39_VACCV	PROTEIN A39.

4.02e-08

4.02e-08	3	109	2.1	327	7	SYFA_ECOLI	PHENYLALANYL-TRNA SYN
2.38e-02	4	107	2.1	2769	7	THYG_BOVIN	THYROGLOBULIN PRECURS
4.70e-02	5	105	2.0	195	7	VC04_VARV	PROTEIN C4.
9.21e-02	6	104	2.0	642	3	HEM1_RAT	5-AMINOLEVULINIC ACID
1.28e-01	7	103	2.0	799	4	ITB5_HUMAN	INTEGRIN BETA-5 SUBUN
1.78e-01	8	98	1.9	191	1	APP2_RAT	AMYLOID-LIKE PROTEIN
8.88e-01	9	97	1.9	197	3	GRPE_ECOLI	HEAT SHOCK PROTEIN GR
1.21e+00	10	100	1.9	203	6	RL15_HUMAN	60S RIBOSOMAL PROTEIN
4.71e-01	11	96	1.9	213	2	CG1_MEDSA	CYCLIN 1 (FRAGMENT).
1.65e+00	12	98	1.9	315	7	VC04_VACCV	PROTEIN C4.
8.88e-01	13	99	1.9	339	2	EBA1_FLAME	ENDO-BETA-N-ACETYLGLU
6.48e-01	14	96	1.9	640	3	HEM1_HUMAN	5-AMINOLEVULINIC ACID
1.65e+00	15	99	1.9	688	7	SYGB_ECOLI	GLYCYL-TRNA SYNTHETAS
6.48e-01	16	97	1.9	700	3	GUNA_BACLA	ENDOGLUCANASE A PRECU
1.21e+00	17	96	1.9	763	1	APP2_HUMAN	AMYLOID-LIKE PROTEIN
1.65e+00	18	97	1.9	875	3	GYRA_ECOLI	DNA GYRASE SUBUNIT A
1.21e+00	19	96	1.9	1133	8	VGLM_SEOU8	M POLYPROTEIN PRECURS
1.65e+00	20	96	1.9	1133	8	VGLM_SEOUS	M POLYPROTEIN PRECURS
1.65e+00	21	96	1.9	1133	8	VGLM_HANTB	M POLYPROTEIN PRECURS
1.65e+00	22	96	1.9	1276	1	BXD_CLOBO	BOTULINUM NEUROTOXIN
1.65e+00	23	96	1.9	1390	5	MET_HUMAN	HEPATOCYTE GROWTH FAC
1.65e+00	24	96	1.9	1451	8	VGL2_CVCAI	E2 GLYCOPROTEIN PRECU
1.65e+00	25	99	1.9	1651	2	CO3_NAJNA	COMPLEMENT C3 PRECURS
6.48e-01	26	91	1.8	221	1	ARP4_TOBAC	AUXIN-INDUCED PROTEIN
7.39e+00	27	93	1.8	261	6	RIPS_PHYAM	ANTIVIRAL PROTEIN S (
4.10e+00	28	92	1.8	316	7	VC04_VACCC	PROTEIN C4.
5.51e+00	29	92	1.8	422	4	K1CR_MOUSE	KERATIN, TYPE I CYTOS

5.51e+00	30	92	1.8	430	1	AROA_STAAU	3-PHOSPHOSHIKIMATE 1-
5.51e+00	31	91	1.8	535	7	UL21_HSV11	PROTEIN UL21.
7.39e+00	32	91	1.8	560	1	60IM_PSEPU	60 KD INNER-MEMBRANE
7.39e+00	33	93	1.8	604	5	NODM_RHIME	GLUCOSAMINE--FRUCTOSE
4.10e+00	34	92	1.8	695	1	APP2_MOUSE	AMYLOID-LIKE PROTEIN
5.51e+00	35	91	1.8	712	6	RS3_CHLRE	PUTATIVE 30 RIBOSOMAL
7.39e+00	36	93	1.8	1071	7	UBPI_YEAST	PUTATIVE UBIQUITIN CA
4.10e+00	37	95	1.8	1157	6	POL_SFV3L	POL POLYPROTEIN (CONT
2.24e+00	38	95	1.8	1161	6	POL_SFV1	POL POLYPROTEIN (CONT
2.24e+00	39	91	1.8	1229	7	SIP3_YEAST	SIP3 PROTEIN.
7.39e+00	40	90	1.7	121	3	GLB_TETTH	MYOGLOBIN (HEMOGLOBIN
9.87e+00	41	90	1.7	253	8	Y29K_NPVAC	HYPOTHETICAL 29.9 KD
9.87e+00	42	90	1.7	274	1	BACH_HALSP	HALORHODOPSIN PRECURS
9.87e+00	43	90	1.7	370	2	DNAJ_ERYRH	DNAJ PROTEIN.
9.87e+00	44	90	1.7	458	2	DESM_XENLA	DESMIN.
9.87e+00	45	90	1.7	798	4	ITB7_HUMAN	INTEGRIN BETA-7 SUBUN
9.87e+00							

ALIGNMENTS

```

RESULT      1
ID  VA39_VACCC      STANDARD;      PRT;      403 AA.
AC  P21062;
DT  01-FEB-1991 (REL. 17, CREATED)
DT  01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT  01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE  PROTEIN A39.
GN  A39R.
OS  VACCINIA VIRUS (STRAIN COPENHAGEN).
OC  VIRIDAE;      DS-DNA      ENVELOPED      VIRUSES;      POXVIRIDAE;
CHORDOPOXVIRINAE;
OC  ORTHOPOXVIRUSES.
RN  [1]
RP  SEQUENCE FROM N.A.
RM  91021027
RA  GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW
    J.P.,

```

RA PAOLETTI E.;
 RL VIROLOGY 179:247-266(1990).
 RN [2]
 RP COMPLETE GENOME.
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
 RA PAOLETTI E.;
 RL VIROLOGY 179:517-563(1990).
 DR EMBL; M35027; PXVACCG.
 DR PIR; E42521; E42521.
 SQ SEQUENCE 403 AA; 45741 MW; 907305 CN;

Query Match 2.8%; Score 144; DB 7; Length 403;
 Best Local Similarity 32.8%; Pred. No. 4.02e-08;
 Matches 22; Conservative 18; Mismatches 21; Indels 6;
 Gaps 5;

D b 1 6 2
 fvdkgdgydk-vyilftdt-igskrivkipy--iaqmclndeggpsslsshrwstflkve 217
 ||: :| :::: :| : : |: | :|::| :| ||| |
 ||::|||
 Q y 2 0 6
 FVN-SVAYGDYIFFFYRETAVEYMNCGKVIYSRVARVCKDDKGGPHQ-SRDRWTSFLKAR 263

Db 218 lecdidg 224
 |:| | |
 Qy 264 LNCSIPG 270

Search completed: Thu May 16 15:34:23 1996
 Job time : 58 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:37:53 1996; MasPar time 2.71
Seconds

373.317 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-66

Description: (1-122) from US08121713B.pep

Perfect Score: 940

S e q u e n c e :

MIYLYTADNVIPKDGLOGAF.....MYSLIVLFQVRIMYLFYEYH 122

1

Scoring table: PAM 150
Gap 11

Searched: 70887 seqs, 8282111 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8 9:part9 10:part10 11:part11 12:part12
13:part13
14:part14

Statistics: Mean 29.350; Variance 116.666; scale 0.252

Pred. No. is the number of results predicted by chance to
have a

score greater than or equal to the score of the result being
printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID	Description	
Pred. No.							
1	940	100.0	122	13	R71385	Variola major virus s	

1.38e-86	2	608	64.7	441	13	R71381	Vaccinia virus semaph
1.63e-51	3	116	12.3	477	13	R74175	Human collapsin.
2.87e-02	4	116	12.3	712	13	R71384	Tribolium semaphorin.
2.87e-02	5	116	12.3	771	13	R71380	Human semaphorin III
2.87e-02	6	109	11.6	724	13	R71383	Drosophila semaphorin
1.14e-01	7	101	10.7	730	13	R71379	Grasshopper semaphori
5.36e-01	8	95	10.1	650	13	R71382	Drosophila semaphorin
1.67e+00	9	85	9.0	325	2	P70428	Polypeptide encoded b
1.06e+01	10	84	8.9	2254	14	R76949	ACCCase.
1.27e+01	11	83	8.8	651	14	R74042	Maize carbonic-anhydr
1.51e+01	12	83	8.8	655	14	R74044	Carbonic-anhydrase.
1.51e+01	13	83	8.8	970	14	R72458	Porphyromonas gingiva
1.51e+01	14	80	8.5	348	7	R38483	Rhodopsin protein.
2.58e+01	15	79	8.4	387	10	R53702	Sequence of castor mi
3.08e+01	16	79	8.4	456	9	R49136	Sequence of lag E whi
3.08e+01	17	78	8.3	510	7	R34674	Insulinoma-associated
3.67e+01	18	78	8.3	2240	11	R67819	Acetyl CoA carboxylas
3.67e+01	19	77	8.2	1206	3	P60679	Sequence of Rift Vall
4.37e+01	20	77	8.2	1206	1	P82995	Rift Valley fever vir
4.37e+01	21	76	8.1	287	10	R53544	Thyroid hormone recep
5.20e+01	22	76	8.1	360	3	P83252	Sequence encoded by t
5.20e+01	23	76	8.1	529	5	R24948	Sequence encoded by h
5.20e+01	24	76	8.1	529	3	R13503	HSF.
5.20e+01	25	76	8.1	764	3	R12504	Canine thyroid stimul
5.20e+01	26	76	8.1	1074	4	R24102	Marek's disease virus
5.20e+01	27	75	8.0	349	1	P90554	Bovine rhodopsin.
6.18e+01	28	74	7.9	293	8	R41268	vWF fragment Arg441-V

7.33e+01	29	74	7.9	293	2	R23659	Mutant Cys-free matur
7.33e+01	30	74	7.9	456	1	P93114	EGR2.
7.33e+01	31	74	7.9	456	11	R63130	Human Egr-2.
7.33e+01	32	74	7.9	470	7	R34467	Encoded by Hepatitis
7.33e+01	33	74	7.9	470	7	R34476	Encoded by Hepatitis
7.33e+01	34	74	7.9	470	7	R34478	Encoded by Hepatitis
7.33e+01	35	73	7.8	59	12	R63135	Egr-1 nuclear localiz
8.70e+01	36	73	7.8	84	1	R04534	Amino terminal portio
8.70e+01	37	73	7.8	89	11	R63136	Egr-1 polynucleotide
8.70e+01	38	73	7.8	250	11	R60153	Nematode-inducible tr
8.70e+01	39	73	7.8	250	7	R33913	Sequence encoded by T
8.70e+01	40	73	7.8	533	1	P93113	Egr-1.
8.70e+01	41	73	7.8	533	11	R63129	Mouse Egr-1 clone OC3
8.70e+01	42	73	7.8	543	1	R24022	Human promyelo-leukae
8.70e+01	43	73	7.8	668	1	P81186	Sequence encoded by n
8.70e+01	44	73	7.8	668	1	P82924	Cowpox virus protein
8.70e+01	45	73	7.8	683	8	R40386	betaIG-H3 protein.
8.70e+01							

ALIGNMENTS

RESULT 1
 ID R71385 standard; Protein; 122 AA.
 AC R71385;
 DT 21-NOV-1995 (first entry)
 DE Variola major virus semaphorin IV protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila;
 Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding
 activity;
 KW modulation; nerve cell growth; immune response; viral
 pathogenesis;
 KW neurological disease; neuro-regeneration; oncological
 infection.
 OS Variola major virus.
 PN WO9507706-A.

PD 23-MAR-1995.
 PF 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 DR WPI; 95-131177/17.
 DR N-PSDB; Q87447.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2; Page 90-91; 101pp; English.
 CC The sequence of the variola major (smallpox) virus semaphorin
 IV protein.
 CC The gene sequence was isolated as the A43R open reading frame
 sequence
 CC from variola based on sequence homology searches of a database
 with the
 CC grasshopper, Tribolium and Drosophila semaphorin sequences.
 The proteins
 CC encoded by the grasshopper semaphorin I (Q87441), human
 semaphorin III
 CC (Q87442), vaccinia virus semaphorin IV (Q87443), Drosophila
 semaphorin I
 CC and II (Q87444-5), Tribolium semaphorin I (Q87446) or variola
 major
 CC (smallpox) virus semaphorin IV genes were used to generate a
 series of
 CC peptides (R70370-R70418), which retain semaphorin receptor
 binding
 CC activity. The semaphorin derived or semaphorin receptor
 derived peptides
 CC are potent modulators of nerve cell growth, immune
 responsiveness and
 CC viral pathogenesis. They can be used in diagnosis and
 treatment of
 CC neurological disease and neuro-regeneration, immune modulation
 and
 CC diagnosis and treatment of viral and oncological infection and
 diseases.

SQ Sequence 122 AA;

Query Match 100.0%; Score 940; DB 13; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1.38e-86;
 Matches 122; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 Search completed: Thu May 16 15:38:03 1996
 Job time : 10 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:37:18 1996; MasPar time 5.36
Seconds

575.185 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-66

Description: (1-122) from US08121713B.pep

Perfect Score: 940

S e q u e n c e :

MIYLYTADNVIPKDGLOGAF.....MYSLIVLFQVRIMYLFYEYH 122

1

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3
7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 40.357; Variance 91.599; scale 0.441

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
Pred. No.							

1	940	100.0	122	6	JQ1845		14R protein - variol
5.14e-146							

2	940	100.0	122	6	H36852	A43R protein - vario
5.14e-146						
3	608	64.7	295	6	JQ1775	SalL9R protein - vac
1.18e-85						
4	608	64.7	403	6	E42521	A39R protein - vacci
1.18e-85						
5	608	64.7	441	6	S29921	hypothetical protein
1.18e-85						
6	117	12.4	120	4	S27251	NADH dehydrogenase (
1.40e-03						
7	117	12.4	120	1	DERZN3	NADH dehydrogenase (
1.40e-03						
8	117	12.4	120	4	S01599	NADH dehydrogenase (
1.40e-03						
9	117	12.4	120	1	DENTN3	NADH dehydrogenase (
1.40e-03						
10	117	12.4	120	1	DELVN3	NADH dehydrogenase (
1.40e-03						
11	117	12.4	120	4	S04434	NADH dehydrogenase (
1.40e-03						
12	117	12.4	772	10	A49069	collapsin - chicken
1.40e-03						
13	116	12.3	711	13	A49423	semaphorin I precurs
1.89e-03						
14	116	12.3	771	11	D49423	semaphorin III precu
1.89e-03						
15	115	12.2	120	4	S04435	NADH dehydrogenase (
2.55e-03						
16	109	11.6	724	10	C49423	semaphorin II precur
1.50e-02						
17	102	10.9	118	9	S53860	NADH dehydrogenase c
1.11e-01						
18	101	10.7	730	10	JH0798	fasciclin IV precurs
1.47e-01						
19	100	10.6	128	3	QQVZF6	F6 protein - vaccini
1.95e-01						
20	100	10.6	128	6	C42513	L5R protein - vaccin
1.95e-01						
21	100	10.6	128	6	B36845	M5R protein - variol
1.95e-01						
22	100	10.6	128	6	S33091	L5R protein - variol
1.95e-01						
23	98	10.4	501	12	S48120	deoxyribodipyrimidin
3.39e-01						
24	98	10.4	587	9	A31776	lactose permease - y
3.39e-01						
25	95	10.1	656	10	B49423	semaphorin I - fruit
7.68e-01						
26	87	9.3	250	9	C60944	cytochrome b - Leish
6.30e+00						
27	87	9.3	324	5	S29619	transforming protein
6.30e+00						
28	87	9.3	324	2	TVRTAS	transforming protein
6.30e+00						

29	85	9.0	833	9	A31593	heat shock transcrip
1.05e+01						
30	85	9.0	833	9	A31592	heat shock transcrip
1.05e+01						
31	85	9.0	4753	10	S27801	LDL receptor-related
1.05e+01						
32	85	9.0	4753	10	A47437	LDL-receptor-related
1.05e+01						
33	84	8.9	118	4	S25944	NADH dehydrogenase (
1.34e+01						
34	84	8.9	302	4	B44349	maturation-promoting
1.34e+01						
35	84	8.9	302	4	A44349	maturation-promoting
1.34e+01						
36	84	8.9	325	2	TVHUAS	transforming protein
1.34e+01						
37	84	8.9	1058	3	WMBE52	UL52 protein - human
1.34e+01						
38	83	8.8	194	11	S28850	membrane protein PMP
1.72e+01						
39	83	8.8	317	8	B41662	probable pheromone-r
1.72e+01						
40	83	8.8	1164	1	RNVZCP	DNA-directed RNA pol
1.72e+01						
41	82	8.7	121	4	C42573	NADH dehydrogenase (
2.20e+01						
42	82	8.7	423	11	B40470	glucocorticoid-induc
2.20e+01						
43	82	8.7	477	6	S25822	transposase (inserti
2.20e+01						
44	82	8.7	491	11	C40470	glucocorticoid-induc
2.20e+01						
45	82	8.7	5126	13	S40450	ryanodine receptor/c
2.20e+01						

ALIGNMENTS

```

RESULT      1
ENTRY       JQ1845      #type complete
TITLE       14R protein - variola major virus
ORGANISM    #formal_name variola major virus
DATE        03-May-1994 #sequence_revision 03-May-1994
#text_change
            03-May-1994
ACCESSIONS  JQ1845
REFERENCE   JQ1832
            #authors    Aguado, B.; Selmes, I.P.; Smith, G.L.
            #journal     J. Gen. Virol. (1992) 73:2887-2902
            #title       Nucleotide sequence of 21.8 kbp of variola major
virus strain
            Harvey and comparison with vaccinia virus.
            #accession   JQ1845
            ##molecule_type DNA

```

##residues 1-122 ##label AGU
##experimental_source strain Harver
SUMMARY #length 122 #molecular-weight 14301 #checksum
8838

Query Match 100.0%; Score 940; DB 6; Length 122;
Best Local Similarity 100.0%; Pred. No. 5.14e-146;
Matches 122; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

D b 1
miylytadnvpkdgllqgafvdkdgtvdkvyilftvtigskrivkipyiaqmclndecgp 60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Q y 1
MIYLYTADNVIPKDGLLQGAFVDKDGTYDKVYILFTVTIGSKRIVKIPYIAQMCLNDECGP 60

D b 6 1
sslsshwrstllkvelecdidgrsysqinhsktikqimiryymyslivlfqvrmylfye 120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Q y 6 1
SSLSSHRWSTLLKVELECDIDGRSYSQINHSTIKQIMIRYYMYSLIVLFQVRIMYLFYE 120

Db 121 yh 122
||
Qy 121 YH 122

RESULT 2
ENTRY H36852 #type complete
TITLE A43R protein - variola virus (strain India-1967)
ORGANISM #formal_name variola virus
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993
#text_change

30-Sep-1993
ACCESSIONS H36852
REFERENCE A36859
#authors Blinov, V.M.
#submission submitted to GenBank, November 1992
#description not shown.
#accession H36852
##status preliminary
##molecule_type DNA
##residues 1-122 ##label BLI
##cross-references GB:X69198

SUMMARY #length 122 #molecular-weight 14301 #checksum
8838

Query Match 100.0%; Score 940; DB 6; Length 122;
Best Local Similarity 100.0%; Pred. No. 5.14e-146;
Matches 122; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

D b 1
miyltadnvipkdglgga fvdkdgtydkvyilftvtigskrivkipyiaqmclndecgp 60

D b 6 1
sslsshrwstllkvelecdidgrsysqinhsktikqimiryymyslivilfqvrmyllye 120

Db	121	yh	122
Qy	121	YH	122

Query Match 64.7%; Score 608; DB 6; Length 295;
Best Local Similarity 91.5%; Pred. No. 1.18e-85;
Matches 86; Conservative 3; Mismatches 4; Indels 1;
Gaps 1;

Db 242 slsshrwstflkvelecdidgrsyrqiihsrtik 275
| | | | | : | | | | | | | | | : | | |

Qy 62 SLSSHRWSTLLKVELECDIDGRSYSQINHSKTIK 95

```

RESULT      4
ENTRY       E42521      #type complete
TITLE       A39R protein - vaccinia virus (strain Copenhagen)
ORGANISM     #formal_name vaccinia virus
             #note       host Homo sapiens (man)
DATE         09-Nov-1990 #sequence_revision 09-Nov-1990
#text_change
             08-Apr-1994
ACCESSIONS   E42521
REFERENCE    A33172
#authors     Johnson, G.P.
#submission  submitted to GenBank, June 1990
#accession   E42521
##status     preliminary
##molecule_type DNA
##residues   1-403 ##label JOH
SUMMARY      #length 403 #molecular-weight 45741 #checksum
8167

```

Query Match 64.7%; Score 608; DB 6; Length 403;
Best Local Similarity 91.5%; Pred. No. 1.18e-85;
Matches 86; Conservative 3; Mismatches 4; Indels 1;
Gaps 1;

```
D               b                     1                      4                        4
y d l y t a d n v i p k d g l r g a f v d k d g t y d k v y i l f t d t i g s k r i v k i p y i a q m c l n d e g g p s   203
      | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | |
Q                                     y                               3
Y-LYTADNVIPKDGLOGAFVDKDGT YDKVYILFTVTIGSKRIVKIPIYAQMCLNDECGPS 61
```

```
Db      204 slsshrwstflkvelecdidgrsyqiihsrtik 237
      |||||:|||||  ||  ||:|||
Qy      62 SLSSHRWSTLLKVELECDIDGRSYSQINHSKTIK 95
```

```

RESULT      5
ENTRY       S29921      #type complete
TITLE       hypothetical protein 15 - vaccinia virus
ORGANISM     #formal_name vaccinia virus
DATE         20-Feb-1995 #sequence_revision 20-Feb-1995
#text_change
            20-Feb-1995
ACCESSIONS   S29921
REFERENCE    S29907
#authors     Amegadzie, B.Y.
#submission  submitted to the EMBL Data Library, January 1991
#accession   S29921
##status     preliminary
##molecule_type DNA
##residues   1-441 ##label AME

```


Query Match 12.4%; Score 117; DB 4; Length 120;
Best Local Similarity 45.8%; Pred. No. 1.40e-03;
Matches 11; Conservative 9; Mismatches 3; Indels 1;
Gaps 1;

Db 57 qfrirrymfalvfvvfdvetvfly 80
|: |||||::|: |:|:| ::|:
Qy 96 QIMIRYYMYSLI-VLFQVRIMYLF 118

Search completed: Thu May 16 15:37:37 1996
Job time : 19 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh,

U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman
algorithm

Run on: Thu May 16 15:36:51 1996; MasPar time 3.34
Seconds

560.446 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-66

Description: (1-122) from US08121713B.pep

Perfect Score: 940

S e q u e n c e :

MIYLYTADNVIPKDGLOGAF.....MYSLIVLFQVRIMYLFYEYH 122

1

Scoring table: PAM 150
Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot31
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6
7:part7
8:part8

Statistics: Mean 42.200; Variance 76.909; scale 0.549

Pred. No. is the number of results predicted by chance to
have a
score greater than or equal to the score of the result being
printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID	Description	
Pred. No.							

1	608	64.7	403	7	VA39_VACCC	PROTEIN A39.	
2.66e-106							
2	608	64.7	441	7	VA39_VACCV	PROTEIN A39.	

2.66e-106	3	117	12.4	120	5	NU3C_MARPO	NADH-PLASTOQUINONE OX
5.26e-05	4	117	12.4	120	5	NU3C_TOBAC	NADH-PLASTOQUINONE OX
5.26e-05	5	117	12.4	120	5	NU3C_MAIZE	NADH-PLASTOQUINONE OX
5.26e-05	6	117	12.4	120	5	NU3C_ORYSA	NADH-PLASTOQUINONE OX
5.26e-05	7	117	12.4	120	5	NU3C_WHEAT	NADH-PLASTOQUINONE OX
5.26e-05	8	115	12.2	120	5	NU3C_SYNY3	NADH-PLASTOQUINONE OX
1.09e-04	9	100	10.6	128	8	VL05_VACCV	PROTEIN L5 (PROTEIN F
1.99e-02	10	100	10.6	128	8	VL05_VARV	PROTEIN L5.
1.99e-02	11	98	10.4	501	5	PHR1_SINAL	DEOXYRIBODIPYRIMIDINE
3.86e-02	12	98	10.4	587	4	LACP_KLULA	LACTOSE PERMEASE.
3.86e-02	13	87	9.3	324	4	MAS_MOUSE	MAS PROTO-ONCOGENE.
1.26e+00	14	87	9.3	324	4	MAS_RAT	MAS PROTO-ONCOGENE.
1.26e+00	15	85	9.0	833	4	HSF_YEAST	HEAT SHOCK FACTOR PRO
2.29e+00	16	84	8.9	118	5	NU3M_MARPO	NADH-UBIQUINONE OXIDO
3.08e+00	17	84	8.9	302	2	CC22_XENLA	CELL DIVISION CONTROL
3.08e+00	18	84	8.9	302	2	CC21_XENLA	CELL DIVISION CONTROL
3.08e+00	19	84	8.9	325	4	MAS_HUMAN	MAS PROTO-ONCOGENE.
3.08e+00	20	84	8.9	1058	7	UL52_HSV11	DNA REPLICATION PROTE
3.08e+00	21	83	8.8	193	5	PMP2_RAT	22 KD PEROXISOMAL MEM
4.13e+00	22	83	8.8	1164	6	RPO2_COWPX	DNA-DIRECTED RNA POLY
4.13e+00	23	82	8.7	121	5	NQO7_PARDE	NADH-UBIQUINONE OXIDO
5.51e+00	24	82	8.7	423	3	GCRC_MOUSE	PROBABLE G PROTEIN-CO
5.51e+00	25	82	8.7	477	7	TRAF_BACTI	TRANSPOSASE FOR INSER
5.51e+00	26	81	8.6	87	2	CYC6_APHFL	CYTOCHROME C6 (SOLUBL
7.35e+00	27	81	8.6	451	4	HSF_XENLA	HEAT SHOCK FACTOR PRO
7.35e+00	28	81	8.6	1164	6	RPO2_VARV	DNA-DIRECTED RNA POLY
7.35e+00	29	80	8.5	143	2	EAG_BACSU	HYPOTHETICAL 16.4 KD

9.77e+00	30	80	8.5	348	5	OPSD_CANFA	RHODOPSIN.
9.77e+00	31	80	8.5	348	5	OPSD_CRIGR	RHODOPSIN.
9.77e+00	32	80	8.5	348	5	OPSD_HUMAN	RHODOPSIN.
9.77e+00	33	80	8.5	348	5	OPSD_SHEEP	RHODOPSIN.
9.77e+00	34	80	8.5	348	5	OPSD_MOUSE	RHODOPSIN.
9.77e+00	35	80	8.5	447	5	NU4M_ANOGA	NADH-UBIQUINONE OXIDO
9.77e+00	36	80	8.5	706	8	YEU3_YEAST	HYPOTHETICAL 81.5 KD
9.77e+00	37	80	8.5	858	2	CHS2_RHIOL	CHITIN SYNTHASE 2 (EC
9.77e+00	38	79	8.4	234	3	GLNA_DUNSA	GLUTAMINE SYNTHETASE
1.29e+01	39	79	8.4	244	5	MLS2_STAAU	RRNA ADENINE N-6-METH
1.29e+01	40	79	8.4	273	6	PSAB_YERPE	CHAPERONE PROTEIN PSA
1.29e+01	41	79	8.4	468	5	NIFB_KLEPN	NIFB PROTEIN.
1.29e+01	42	79	8.4	593	7	VG13_BPML5	GENE 13 PROTEIN (GP13
1.29e+01	43	79	8.4	649	7	TFC1_YEAST	TRANSCRIPTION FACTOR
1.29e+01	44	79	8.4	739	2	DCLY_HAFAL	LYSINE DECARBOXYLASE
1.29e+01	45	79	8.4	1164	6	RPO2_VACCV	DNA-DIRECTED RNA POLY
1.29e+01							

ALIGNMENTS

RESULT 1

ID VA39 VACCC STANDARD; PRT; 403 AA.

AC P21062;

DT 01-FEB-1991 (REL. 17, CREATED)

DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)

DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)

DE PROTEIN A39.

GN A39R.

OS VACCINIA VIRUS (STRAIN COPENHAGEN).

OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;

OC ORTHOPOXVIRUSES.

RN [1]

RP SEQUENCE FROM N.A.

RM 91021027

RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,

